

8-2010

IMPROVING QUANTITATIVE TREATMENT RESPONSE MONITORING WITH DEFORMABLE IMAGE REGISTRATION

Blake A. Cannon

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Medical Biomathematics and Biometrics Commons](#), and the [Neoplasms Commons](#)

Recommended Citation

Cannon, Blake A., "IMPROVING QUANTITATIVE TREATMENT RESPONSE MONITORING WITH DEFORMABLE IMAGE REGISTRATION" (2010). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 80.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/80

This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

IMPROVING QUANTITATIVE TREATMENT RESPONSE MONITORING WITH
DEFORMABLE IMAGE REGISTRATION

by

Blake Cannon, M.S.

APPROVED:

Lei Dong, Ph.D., Supervisory Professor

Mary Martel, Ph.D.

Osama Mawlawi, Ph.D.

Susan Tucker, Ph.D.

David Schwartz, M.D.

APPROVED:

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

IMPROVING QUANTITATIVE TREATMENT RESPONSE MONITORING WITH
DEFORMABLE IMAGE REGISTRATION

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
M. D. Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Blake Cannon, M.S.
Houston, Texas

August 2010

*For Cara and Avery,
whose sacrifice, support, patience, and love made this possible*

Acknowledgements

I would like to thank Dr. Dong for his constantly positive support and mentorship. His patience and work ethic were a constant source of inspiration. I would like to also thank my committee for dedicating so much of their time to make this work what it is today: to Dr. Schwartz for providing the cohort data and always being an enthusiastic advocate, to Dr. Tucker for her insightful conversations on data analysis, and to Dr. Mawlawi for pushing me to better communicate my work to others. Finally, I would like to thank Dr. Martel for serving on my candidacy exam committee and her counsel and support navigating the waters of professional medical physics.

Abstract

Quantitative imaging with ^{18}F -FDG PET/CT has the potential to provide an in vivo assessment of response to radiotherapy (RT). However, comparing tissue tracer uptake in longitudinal studies is often confounded by variations in patient setup and potential treatment induced gross anatomic changes. These variations make true response monitoring for the same anatomic volume a challenge, not only for tumors, but also for normal organs-at-risk (OAR). The central hypothesis of this study is that more accurate image registration will lead to improved quantitation of tissue response to RT with ^{18}F -FDG PET/CT. Employing an in-house developed “demons” based deformable image registration algorithm, pre-RT tumor and parotid gland volumes can be more accurately mapped to serial functional images. To test the hypothesis, specific aim 1 was designed to analyze whether deformably mapping tumor volumes rather than aligning to bony structures leads to superior tumor response assessment. We found that deformable mapping of the most metabolically avid regions improved response prediction ($P < 0.05$). The positive predictive power for residual disease was 63% compared to 50% for contrast enhanced post-RT CT. Specific aim 2 was designed to use parotid gland standardized uptake value (SUV) as an objective imaging biomarker for salivary toxicity. We found that relative change in parotid gland SUV correlated strongly with salivary toxicity as defined by the RTOG/EORTC late effects analytic scale (Spearman's $\rho = -0.96$, $P < 0.01$). Finally, the goal of specific aim 3 was to create a phenomenological dose-SUV response model for the human parotid glands.

Utilizing only baseline metabolic function and the planned dose distribution, predicting parotid SUV change or salivary toxicity, based upon specific aim 2, became possible. We found that the predicted and observed parotid SUV relative changes were significantly correlated (Spearman's $\rho = 0.94$, $P < 0.01$). The application of deformable image registration to quantitative treatment response monitoring with ^{18}F -FDG PET/CT could have a profound impact on patient management. Accurate and early identification of residual disease may allow for more timely intervention, while the ability to quantify and predict toxicity of normal OAR might permit individualized refinement of radiation treatment plan designs.

Table of Contents

Acknowledgements	iv
Abstract	v
Chapter 1	1
Introduction	1
Functional Response Monitoring	2
Effects of Anatomic Deformation and Patient Set-up Variation	4
Tumor Response	6
Normal Tissue Response	9
Central Hypothesis	11
Specific aim 1	11
Specific aim 2.a	11
Specific aim 2.b	12
Chapter 2	13
Radiotherapy and the Head and Neck	13
Cancers of the Head and Neck	13
Treatment of Head & Neck Cancers	13
The Human Salivary System	15
Parotid Radiosensitivity	17
Assessing Salivary Toxicity	20
Chapter 3	23
Specific Aim 1	23
Purpose:	23

Methods and Materials:	24
Patient Cohort.....	24
RT and Systemic Treatment	25
Functional Imaging	25
Anatomic Imaging	26
Post-Treatment Surveillance	27
Pathologic Tissue Assessment.....	27
Tumor Volume Delineation	28
Image Registration.....	30
Volume Alignment Workflow	32
Calculating SUV in the CAT Workspace.....	34
Receiver Operating Characteristic Analysis.....	35
Results:.....	37
Patient Characteristics.....	37
SUV and Treatment Response	39
Predicting RT Outcomes with SUV	48
Response	53
Predicting RT Outcomes with SUV Derived From Bony Aligned Volumes	60
Comparing SUV Derived from Bony and Deformable Alignment.....	64
Discussion:	65
Conclusions:	68
Chapter 4	70
Specific Aim 2a	70

Purpose	70
Methods and Materials	72
Patient Cohort.....	72
Sialometric Evaluation	72
Subjective Evaluation	74
Xerostomia Grade.....	74
Volume Delineation.....	76
Results:.....	79
Patient Characteristics	80
Baseline Saliva Output	80
Correlating Baseline Salivary Function and SUV.....	86
Post-RT Salivary Output	90
Correlating Post-RT Salivary Output and SUV	94
Discussion:	101
Conclusions:	105
Chapter 5	106
Specific Aim 2b	106
Purpose	106
Methods and Materials:	107
Patient Cohort.....	107
Phenomenological Dose-Metabolic Response Model.....	108
Metabolic Dose	109
Volume Delineation.....	110

Volume Delineation.....	111
Co-registration of Planned Dose and SUV	111
Validation of the Models	113
Results:.....	114
Fractional SUV and Metabolic Dose	114
Validation of the Mean Model	118
Voxel-based Fractional SUV and Metabolic Dose	123
Validation of the Voxel Model	126
Discussion	131
Discussion	132
Conclusions	138
Chapter 6	139
Summary and Conclusions	139
Central Hypothesis:	139
Conclusions	139
Appendix	143
Addhdrinfo	143
Scanorder	144
Pthdrinfo	147
PETThreshold.....	149
petctfusion	150
OpenPinnacleImages	151
OverlayImg	152

SoftTissue.....	153
RTalpha	153
DualMap	154
Metabolic Dose	156
References	166
Vita	190

List of Illustrations

Figure 1.1 Extreme anatomic and setup variation.....	7
Figure 3.1 Post Treatment Surveillance Workflow	38
SUV and Treatment Response	39
Figure 3.3 SUV _{Max} histograms	42
Figure 3.4 Deformably aligned PET-based threshold contour	44
Figure 3.5 Deformably aligned GTV contour.....	46
Figure 3.6 ROC curves for the deformably aligned volumes.....	51
Figure 3.7 Bony and deformed SUV _{Peak} distributions.....	55
Figure 3.8 Bony and deformed PET threshold contour	57
Figure 3.9 Bony and deformed GTV	58
Figure 3.10 ROC curves for rigidly aligned volumes	62
Figure 4.1 Parotid anatomical boundaries	78
Figure 4.2.....	79
Figure 4.3 Baseline salivary mass	83
Figure 4.4 Baseline salivary mass	84
Figure 4.5 Patient perceived baseline salivary function	85
Figure 4.6 Baseline parotid SUV distributions.....	87
Figure 4.7 Stimulated saliva mass versus SUV	88
Figure 4.8 Baseline XQL score versus parotid SUV	89
Figure 4.9 Histogram distribution of post-RT unstimulated salivary mass.....	91
Figure 4.10 Histogram distribution of post-RT stimulated salivary mass.....	92
Figure 4.11 Histogram distribution of fractional stimulated saliva mass.....	93

Figure 4.12 Unstimulated saliva mass and SUV	95
Figure 4.13 Stimulated saliva mass and SUV	96
Figure 4.14 Post-RT XQL score and SUV	98
Figure 4.15 Fractional stimulated saliva mass and SUV	99
Figure 4.16 Fractional SUV and xerostomia grade	100
Figure 5.1 Metabolic Dose	110
Figure 5.2 Fractional SUV dose response	116
Figure 5.3 Fractional SUV and D_{Met}	117
Figure 5.4 Fractional SUV and D_{Met}	118
Figure 5.5 Fitting results for the mean model.....	121
Figure 5.6 Mean model validation	122
Figure 5.7 Predicted SUV and observed saliva changes	123
Figure 5.8 Voxel-based dose-SUV response.....	125
Figure 5.9 Fitting results for the voxel model	129
Figure 5.10 Voxel model validation	130
Figure 5.11 Predicted SUV and observed saliva changes	131
Figure 5.12 Parotid DVH comparisons.....	137

List of Tables

Table 3.1 Average SUV values derived from deformably mapped volumes for responding and non-responding patients.....47

Table 3.2 Accuracy of response prediction with deformably mapped volumes..52

Table 3.3 Average SUV values for responding and non-responding patients from rigidly aligned volumes.....59

Table 3.4 Accuracy of response prediction with rigidly aligned volumes63

Table 4.1 RTOG/EORTC late effects analytic scale75

"It was an act of desperation. For six years I had struggled with the blackbody theory. I knew the problem was fundamental and I knew the answer. I had to find a theoretical explanation at any cost..."

-Max Planck

Chapter 1

Introduction

Implementation of anatomic imaging modalities, like computed tomography (CT), has found extensive use in the oncologic setting. From initial staging, to treatment planning, to assessing treatment response, CT is ubiquitous [1-4]. In 2000, Therase et al. presented the results of a multi-institutional collaborative effort to evaluate the existing criteria for treatment response in solid tumors set forth by the World Health Organization (WHO) in the 1970s [5]. The new guidelines specified the preferred use of imaging techniques over clinical evaluation to establish response in solid tumors based upon visible reduction in the size of the tumor. The criteria established that for measurable tumors ≥ 10 mm in at least one dimension, the complete disappearance of all target lesions is a complete response, a 30% reduction in the longest diameter of target lesions is a partial response, and a 20% increase in the longest diameter of the target lesion is deemed progressive disease. Stable disease is defined as having insufficient shrinkage or enlargement to classify as partial response or progressive disease, respectively.

However, lesion measurement with anatomic imaging modalities has been shown to be a poor prognostic tool for response to therapy [6-11]. In one study by Jones et al., 56% of patients had incorrect CT-based T-staging when

compared to surgically resected specimens [3]. For regional metastases, decreases in the largest axial dimensions of lymph nodes calculated from CT images were not found to be a significant predictor of positive surgical specimens [7]. These difficulties are not limited to CT imaging. While magnetic resonance imaging (MRI) may have superior soft tissue contrast, previous studies have demonstrated similarly poor results in predicting tumor response [6]. One of the major obstacles in predicting response to treatment with anatomic modalities is differentiating residual disease from necrosis or surrounding normal tissue. However, it is expected that imaging functional change in target tissues, which often precedes visible anatomic change, may improve outcome prediction [12, 13].

Functional Response Monitoring

In recent years, positron emission tomography (PET) using the radiolabeled glucose analog, Fluorine-18-2-fluoro-2-deoxy-D-glucose (FDG), has gained favor in the oncology setting [14-18]. Exploiting an increased reliance on glycolysis for energy production, the preferential uptake of ^{18}F -FDG by malignant tissue provides excellent tumor to background contrast. Multiple studies have demonstrated the superiority of PET in the staging of cancers over anatomic based modalities [19]. Sensitivities and specificities found in the literature range from 84% to 86% and 88% to 93%, respectively. With the development of integrated PET/CT systems, inherently co-registered functional and anatomic images have further improved the usefulness of this modality [9].

Veit-Haibach et al. assessed the accuracy of PET, CT, and PET/CT in staging head and neck cancers and found integrated PET/CT to be significantly more accurate in overall TNM staging [20].

The concentration of radiolabeled tracers can be assessed in two general ways: qualitative and quantitative. In qualitative assessment the experienced nuclear medicine physician compares the intensity and distribution of FDG to normal surrounding tissues, like the liver, in order to classify patients as being PET positive or negative. However, this method can be limited by equivocal FDG-uptake[21].

The second method of assessment is quantitative. Uptake of FDG into cancer cells is directly related to the number of viable cells [22]. Therefore, calculation of uptake reduction should be a surrogate of tumor cell killing. The net rate of ^{18}F -FDG trapped within the cell can be calculated through graphical analysis using the Patlak-Gjedde plot. Utilizing the following equation,

$$c(t) = \lambda * c_p(t) + K_i \int_0^T c_p(t) dt$$

where, $c(t)$ is the activity concentration measured by the imaging system at time t , $c_p(t)$ is the activity concentration in the plasma, K_i is the rate of activity transfer into tissue, and $\lambda * c_p(t)$ is non-trapped FDG in tissue a plot of the area under the time activity curve versus the activity concentration in the tissue with a slope equal to K_i can be found with simple linear regression analysis [23]. This

analysis is often simplified and activity concentration can be quantified with the standardized uptake value (SUV) [20]. The SUV is a simplified version of the equation used in the Patlak-Gjedde analysis. The basic assumption is that if the imaging time point is sufficiently far from the injection time point the activity concentration will be linearly correlated with FDG trapped in the tissue. The time activity curve is assumed to be proportional to the injected dose normalized by some factor describing the distribution of activity in the patient. One common normalization factor is patient mass [24]. The concentration of FDG in the plasma at this late time point is expected to be very low, i.e., $\lambda^*c_p(t)$ is assumed to be zero. The SUV can then be computed using the following relation,

$$K_i \propto \frac{Q_i [MBq / mL]}{\frac{Q_{injected} [MBq]}{W [Kg]}} = SUV_{BW}$$

where Q_i represents the concentration of radiotracer in the tumor or tissue of interest, $Q_{injected}$ is the injected activity, and W is the patient weight in kilograms [24]. The SUV has been shown to correlate well with more intensive methods of calculating tissue activity concentration [25].

Effects of Anatomic Deformation and Patient Set-up Variation

Comparing the same anatomic volume on serial imaging studies is challenging due to two main factors: treatment induced anatomic changes and

patient setup variation. Considering the former, multiple studies have demonstrated the potential anatomic variation in head and neck cancer patients resulting from treatment induced tumor and nodal contraction and weight loss. Barker et al. utilized daily CT imaging of head and neck cancer patients with an integrated CT-linear accelerator system and found the gross tumor volume (GTV) decreases at a median rate of 1.8% per treatment day culminating in a total loss of 70% of the initial volume. The center of mass had shifted position asymmetrically by a median of 3.3 mm. For the parotid glands, the median volume was found to decrease 0.19 cm^3 per day with a median medial shift of 3.1 mm that correlated with weight loss. In figure 1.1, an extreme example of treatment induced anatomic change and patient setup variation is illustrated. Previous response monitoring studies have largely ignored the effects of anatomic and patient setup variation. By utilizing SUV_{MAX} , the maximum pixel value in an FDG-avid region, to sample tumor burden, the vast majority of investigators were able to avoid the need to accurately map baseline volumes to post-RT imaging studies. However, SUV_{MAX} is not an optimal metric because of its strong dependence on noise. In studies where region based analysis was employed, the solution to anatomic and setup variation is usually rigid registration. However, as depicted in figure 1.1 d), rigid alignment of the GTV and parotid contours on post-RT studies appears to be entirely inadequate. The rigidly aligned contours extend beyond the patient's anatomy and contain undesired tissue types, as can be seen for the right parotid contour in figure 1 d). The main goal of this project is to improve quantitation in treatment response

monitoring with ^{18}F -FDG PET/CT. To achieve this goal we will combine advanced image registration techniques with RT treatment designs and serial functional images.

Tumor Response

The potential of PET/CT to improve treatment response monitoring in cancer patients has been noted for some time. In one early study from 1994, Greven et al. found that 6 of 7 head and neck cancer patients with increasing FDG accumulation following RT were positive for persistent disease, while all responding patients were found to have decreases in tracer uptake [26]. Since this initial study multiple reports have investigated ^{18}F -FDG PET/CT to determine its effectiveness in response monitoring [19]. Studies by Allal et al., Kitagawa et al., and Kao et al. found that SUV_{Mean} performed well as a predictor for local control in head and neck cancers [27-29]. One recent study by Yao et al. demonstrated the high negative predictive value of SUV in response monitoring studies (roughly 99%) [30]. In contrast, studies by Arslan et al., Yen et al., and a follow-up by Greven et al. found quantitative measurement of FDG-uptake not to be useful in response monitoring [31-33].

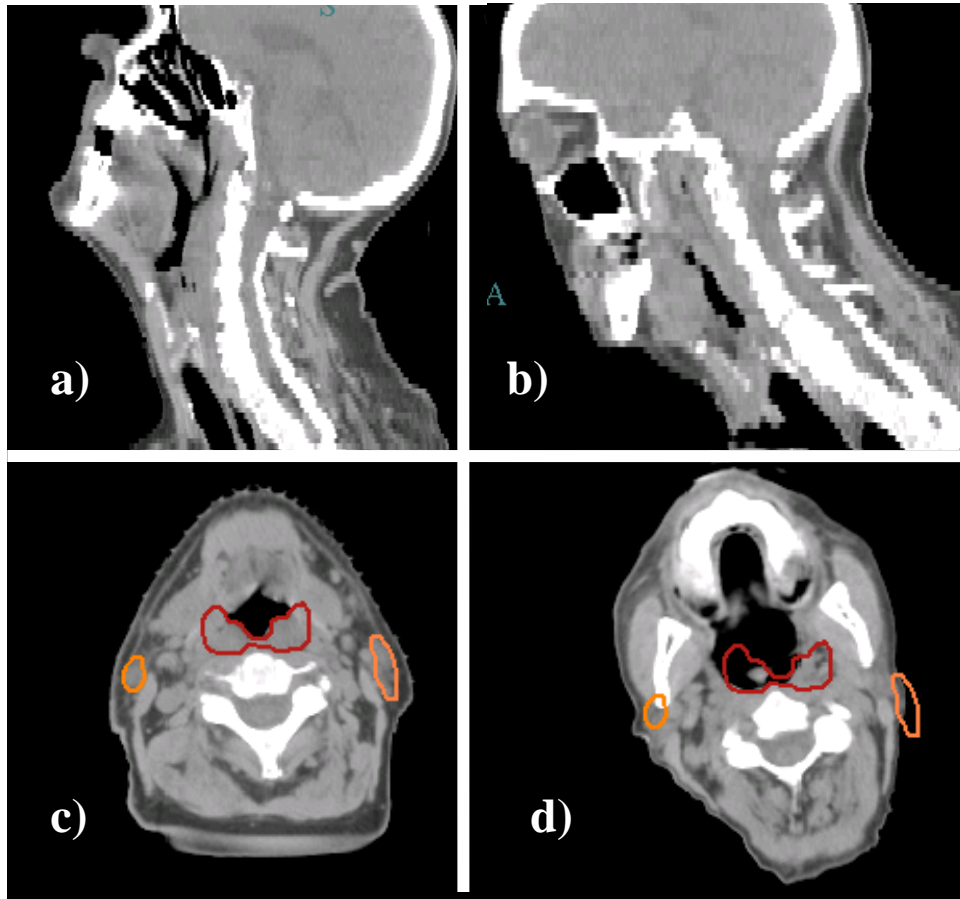


Figure 1.1 Extreme anatomic and setup variation

An extreme example of setup and anatomic variation possible in response monitoring studies between pre-RT CT (a and c) and the post-RT CT (b and d) imaging studies. Axial views are presented with parotid (orange) and gross tumor volume contours (maroon) as originally contoured on the planning CT images and rigidly aligned to the C2 vertebra.

A recent study by Moeller et al. found that while SUV was significantly different between responding and non-responding patients, it offered no improvement over conventional CT in positive (58% vs. 50%) and negative predictive values (96% vs. 97%) and little improvement in positive predictive value (57% vs. 50%) [34]. These examples illustrate that quantitative treatment response monitoring with PET/CT has yet to live up to its initial promise.

While many studies have supported the overall philosophy behind PET imaging of tumor response, namely that reduced uptake of ^{18}F -FDG generally corresponds to better prognosis for patients, questions still remain. Consider the issue of sampling FDG-uptake on PET images. By far the most popular metric to assess tumors has been SUV_{Max} . The justification frequently cited for its use is that the term represents the maximum tumor burden. However, SUV_{Max} is known to be strongly affected by image noise. In studies by Boellard et al. and Falen et al., SUV_{Max} variations in excess of 50% were observed [35, 36]. Alternatively, region-of-interest (ROI) based analysis may be used to determine an average tissue uptake. With volume based techniques, the question then arises of whether automated or manually defined regions should be utilized. More recently, SUV_{Peak} has been suggested as metric to assess tumor viability. SUV_{Peak} is defined as an approximately 1 cm^3 volume centered on the maximum pixel value [37]. In light of expected anatomic changes discussed earlier, recommendations to compare the same volume on serial imaging studies present profound challenges [38].

Normal Tissue Response

Treatment of head and neck cancer often results in the salivary glands receiving high mean doses of radiation (28 to 45 Gy) [39]. Irradiation of the parotids can lead to reduced salivary flow and the subjective assessment of dry mouth, called xerostomia [40, 41]. Xerostomia is the primary morbidity associated with RT for head and neck cancers with 65% of patients reporting severe to moderate xerostomia after one year [40]. Saliva produced by the major and minor salivary glands plays a vital role in mastication, speaking, and the mucosal immune system with the parotid glands producing the majority of the salivary volume (approximately 60%). Salivary flow reduction can affect not only quality of life (QOL), but lead to oral fissures, infection, and potential malnutrition [40]. Salivary function is typically evaluated with patient self-reported QOL questionnaires to gauge the severity of xerostomia, or measurement of saliva volume [42]. However, these techniques often perform poorly and have large variations.

Imaging salivary gland function offers a non-invasive method to assess radiation induced damage. Salivary gland scintigraphy (SGS) is one method that has been utilized to image salivary function by measuring the concentration of ^{99m}Tc -pertechnetate remaining in the glands following the administration of a salivation inducing agent [43]. Previous reports assessing function after RT have found good correlations between activity concentrations measured with scintigraphy and saliva volume [44]. Recently, SGS has been used to perform

dose-response analysis in the parotid glands [45, 46]. However, these techniques were limited to 2D planar images of the salivary glands without any inherently aligned anatomic images. Given the parotid gland shrinkage and medial shifting discussed above, automatic alignment of pre-RT volumes is desirable. With the increasing use of ^{18}F -FDG PET/CT in initial staging and follow-up for head and neck cancers, incidental collection of parotid uptake may prove to be a readily accessible alternative to additional nuclear medicine studies or salivary flow measurements.

Central Hypothesis

We hypothesize that more accurate image registration will lead to improved quantitation of tissue response to radiotherapy with ^{18}F -FDG PET/CT imaging.

Specific aim 1

To apply deformable image registration techniques to images acquired during a prospective clinical trial designed to assess and predict RT outcomes with ^{18}F -FDG PET/CT.

Working Hypothesis: Voxel-by-voxel mapping of primary tumor volumes to serial ^{18}F -FDG PET/CT scans will significantly improve region based quantitative tumor response analysis.

Specific aim 2.a

Evaluate the potential to determine normal organ-at-risk (OAR) toxicity in patients treated with intensity modulated radiotherapy (IMRT) for head and neck cancers with ^{18}F -FDG PET/CT.

Working Hypothesis: Relative change in parotid gland uptake of ^{18}F -FDG will function as an objective imaging biomarker of oral complication following RT.

Specific aim 2.b

To quantify and model the planned-dose-functional-response relationship of the parotid glands in head and neck cancer patients treated with IMRT.

Working Hypothesis: A predicted imaging biomarker for oral complication following RT will correlate closely with whole mouth stimulated salivary flow.

Chapter 2

Radiotherapy and the Head and Neck

Cancers of the Head and Neck

Head and neck cancers make up roughly 6% of all cancers diagnosed in the United States. This corresponds to 46,000 patients suffering from a cancer originating in the head and neck region each year. An estimated 11,000 individuals will die from this disease in the US. Worldwide, over 600,000 patients will be diagnosed with 350,000 deaths. The majority of these cancers arise in the epithelial cells that form the protective linings of cavities and are known as squamous cell carcinomas [47]. This makes alcohol and tobacco consumption significant risk factors for this histology [48]. Additional risk factors include human papilloma virus (HPV) infection, poor oral hygiene, wood dust inhalation, and asbestos to name a few. The primary disease sites of the head and neck regions are the nasopharynx, oropharynx, hypopharynx, and larynx. Among males, cancers of the pyriform sinus, in the hypopharynx, contribute the most patients per 100,000 (2.33) in the US. In females, cancers at this site also supply the most patients (1.67).

Treatment of Head & Neck Cancers

The majority of patients with head and neck cancers present with advanced disease (Stage III and IV), with only about 1/3 presenting with early

stage disease [48]. Although staging varies between sites, primary advanced-stage tumors in the head and neck are typically approximately 4 cm in size as defined by the American Joint Commission on Cancer (AJCC). Typical treatment strategies for late stage disease are surgery, radiation, and chemotherapy. Progression free survival rates seen for all cancers of the oral cavity, oropharynx, hypopharynx, and larynx are 36% at five years with overall survival rates of approximately 40% [49-51]. The concurrent delivery of chemotherapeutic agents like cisplatin can improve 5-year disease free survival to 47% with overall survival rates of 53% [50]. Intensity modulated radiation therapy (IMRT) has shown promising results in terms of locoregional control 94%, but acute toxicity results have not been favorable [52, 53]. Typical treatments for head and neck cancers consist of cumulative doses of 63 – 70 Gy given in 30 – 35 fractions of 1.8 – 2.2 Gy 5 days/week. Alternative fractionations schedules have also led to improvements in 5-year disease free survival. Hyperfractionation regimens have been shown to offer 8% absolute benefit. Phase III clinical trials have led to recurrence rates within the first two years of at least 50% [54, 55]. Treatment following failure is often salvage surgery or systematic therapy with methotrexate [56]. Another option that has led to improvement in progression free survival is re-irradiation. With the recurrence rates seen in this patient population early identification of residual disease is critical for patient management and the use of additional therapies.

Because of the complex anatomy in the head and neck region normal tissue complications are common. Loss of taste, dysphagia, dental carries, and

hearing loss are common symptoms [57]. One of the more frequent morbidities is radiation-induced xerostomia. Up to 65% of patients will report oral dryness one year after treatment [40]. Prevention strategies do exist, but the high doses delivered make damage to the salivary glands a significant issue.

The Human Salivary System

Saliva is a vital body fluid. It performs crucial roles in mastication, digestion, swallowing, oral fauna regulation, and speech. Saliva is composed of approximately 99% water. Other constituents include, but are not limited to, lipids, amino acids, and proteins. The protein components consist primarily of α -amylase (30%) which begins the digestion of polysaccharides. Additionally, saliva provides an excretory route for metabolic waste products like urea. The secretory function of the salivary glands is controlled by the autonomic nervous system. Saliva production can be stimulated through taste, smell, chewing, or exogenous stimulants [41].

Saliva is produced primarily in three paired organs in the upper aerodigestive tract. However, there are multiple minor salivary glands distributed throughout the buccal cavity. The number and arrangement will vary among patients. The three major paired glands are: the parotid, submandibular, and sublingual glands. The major salivary glands produce over a liter of saliva per day [58]. The parotid glands can produce anywhere from roughly half of the salivary volume up to 70% of the total volume. However, saliva produced in the parotid glands is primarily created during stimulation. In unstimulated

conditions, the contribution is much smaller with the submandibular glands taking the dominant role, producing as much as 80% of the saliva volume. The sublingual glands are typically responsible for 2 -5% of the saliva produced [59-61].

The major salivary glands have a well defined structure. The organs consist of ducts, myoepithelial cells, connective tissue, and the saliva producing unit, the acini. The acinar cells may be either mucous or serous exocrine cells. The acinar cells are contained in many spherical clusters throughout the glands. Radiating across the acinar cell, the processes of myoepithelial cells embrace the acinar cell and contract to force saliva into the collecting ducts. The myoepithelial cells comprise about 1.6% of the volume of the parotid glands [60]. Saliva moves through the secondary ductile system and into the intraglandular main duct. In the case of the parotid glands, the saliva then traverses the extraglandular segment of Stensen's duct which is superficial to the masseter muscle before entering the buccal cavity through the orifice of Stensen's duct. The parotid glands contain serous acinar cells, while the submandibular and sublingual glands contain both serous and mucous acinar cells. Qualitatively, the saliva produced in the submandibular and sublingual glands is thick and sticky, with the parotids producing a watery product.

Salivary function can be diminished by many factors e.g., pathologic sources like Sjogren's syndrome or through damage induced by ionizing radiation. Patients receiving radiation therapy for head and neck cancers will

routinely experience high doses to the salivary glands. However, decreases in salivary flow are detectable in as little as 24 hours after doses as small as 2 Gy [62]. Within the first week of treatment salivary flow rates can drop by 50% or more [63-65]. Salivary flow rates will continue to decline throughout the treatment course. In one study of IMRT designed to spare the salivary glands, the contralateral and ipsilateral parotid glands received mean doses of 22 Gy and 55 Gy, respectively, whereas mean doses to the ipsilateral submandibular gland were on the order of 67 Gy. Contralateral submandibular glands receive mean doses of 58 Gy [42]. No appreciable salivary output was measured for the ipsilateral glands. On the other hand, measured salivary flow for contralateral parotid glands decreased initially, but approached pre-RT levels at one year post-RT [42].

Parotid Radiosensitivity

The response of salivary gland constituent tissue to ionizing radiation is complex and not well understood. The saliva producing units of the salivary glands, acinar cells, are the assumed targets in radiation damage leading to reduced salivary flow. The acinar cells are reverting postmitotic cells. Cells of this type will exhibit little regular division and variable differentiation. Consequently, these targets are expected to be relatively less radiosensitive than lymphocytes, for instance, and exhibit mostly late effects.

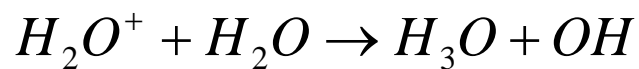
However, as discussed above reduced salivary flow can be detected in as little as 24 hours. One classic explanation is that radiation induced apoptosis is

responsible for the high acute radiosensitivity of the acinar cells, yet studies in rhesus monkeys failed to identify appreciable levels of apoptosis [66, 67]. Additionally, similarly low levels of apoptosis, less than 2%, have been observed in rat acinar cells [68]. Further complicating the picture, Roesink et al., using salivary gland scintigraphy, observed that while uptake of radiotracer was not diminished in the salivary glands, salivary flow was reduced.

To account for the complicated radiation response observed in the salivary glands, hybrid models of radiosensitivity have been proposed. In one model proposed by Konings et al., the acute radiosensitivity was proposed to result from radiation damage to the plasma membrane of the secretory cells [69]. The proposed model is similar to the indirect action of radiation model that results in damage to nuclear DNA. Briefly, after interaction with a photon, a water molecule becomes ionized. The chemical reaction is,



where, H_2O^+ is the ion radical. The ion radical has a short lifetime and decays to form a free radical, which has an unpaired electron. The chemical reaction to form the hydroxyl radical, OH is,



The highly reactive hydroxyl radical is responsible for much of the damage to nuclear DNA. Damage to the cellular membrane is initiated by a process known as lipid peroxidation. In this reaction, the oxygen radical targets the phospholipids that make up the cell membrane. After removing hydrogen from the fatty acid chain, a lipid alkyl radical will be formed. The lipid alkyl radical then reacts with oxygen to form a peroxy radical. The peroxy radical then removes a hydroxyl radical from the neighboring fatty acid chain, producing new alkyl radicals and inducing a chain reaction. One of the many effects of peroxidation of cellular membrane lipids is the disruption of membrane receptors. The disruption of the cell signaling pathway for saliva secretion is hypothesized to be the source of the acute response of the parotid glands to ionizing radiation [70].

The mechanism for late damage is the lack of cell replacement due to the loss of progenitor cells. The progenitor cells in the parotid acini are located at the distal segments of the ductile system. This classic model of late response is supported by studies in the rat parotid [71]. Following irradiation with 30 Gy, the parotid glandular connective tissue increased from 10% to 60% of the volume. The relative number of acinar cells decreased from 80% to 20% of the volume. Following these histopathological findings, the salivary flow was reduced by 90%.

Extensive studies have been carried out to assess the functional outcome the parotid glands after RT. In one study from Eisbruch et al., salivary flow

measurements were collected from 88 patients at multiple times points up to 1 year post-RT [42]. Parotid gland response was best described by dose thresholds. Significantly higher unstimulated salivary flow was seen for glands receiving a mean whole organ dose less than 24 Gy. For stimulated flow, parotid glands receiving a mean dose less than 26 Gy had significantly better function. Chao et al. observed that the parotid glands lose function exponentially with dose. Specifically, the parotid glands were found to lose 5.4% of the stimulated salivary flow at 6 months for every Gy delivered [39]. In contrast, Roesink et al. observed parotid gland functional response to be best described as linear with no threshold dose [45]. While the exact functional form of dose response for parotid gland functional response remains unknown, keeping the mean dose to the parotid glands less than or equal to 26 Gy has been shown to be a significant predictor of salivary toxicity. The $TD_{50/5}$ for the parotid glands is 46 Gy [72].

Assessing Salivary Toxicity

Assessing salivary toxicity is accomplished primarily through two methods: objective and subjective. Objective evaluation may take the form of physician assessment based upon the Radiation Therapy and Oncology Group (RTOG) and European Organization for Research and Treatment of Cancer (EORTC) SOMA scale [73, 74]. Graded 1 – 4, representing increasing severity of toxicity, the scale is meant to capture the response of individual organs. Another method of objective assessment is measurement of the saliva produced

by each of the major glands or from the whole mouth with saliva mixed. Based upon relative changes in the quantity of saliva produced, the EORTC/RTOG scale can be utilized to also assign a grade, 1 – 4, to the salivary toxicity. Subjective evaluation is primarily carried out through patient self-reported quality of life questionnaires [75]. The patient answers a series of questions relating his or her difficulty at performing common tasks like speaking or eating as a result of dryness. Each question is ranked on a 1 – 10 Likert scale and a final summary xerostomia quality of life (XQL) score can be calculated for the questionnaire.

Each metric has limitations. Because xerostomia is the perception of dry mouth, patient self-reported scores may best reflect the symptoms experienced by the patient, but these are fundamentally limited by their subjective nature. On the other hand, XQL scores have been shown to correlate significantly with measured salivary flow rates [76]. In the case of physician assessed toxicity, salivary flow rates were not found to correlate with xerostomia grade [76]. Interestingly, physician assessment was found to underestimate the severity of toxicity when compared to patient self-reported scores. Finally, for measurements of salivary flow, large normal ranges are often found. In one study by Ship et al., untreated individuals with the lowest and highest 10th percentiles of produced saliva mass had similar oral health [77]. Further complicating the use of salivary flow are large intra-patient variations that have been observed on the order of 50%. The poor performance of these

cumbersome additional procedures coupled with ambiguous endpoints presents an opportunity for novel methods to assess salivary toxicity.

Chapter 3

Specific Aim 1

To apply deformable image registration techniques to images acquired during a clinical trial designed to assess and predict RT outcomes with ^{18}F -FDG PET/CT.

Working Hypothesis: Voxel-by-voxel mapping of primary tumor volumes to serial ^{18}F -FDG PET/CT images will improve quantitative tumor response analysis.

Purpose:

Assessing and predicting the response of tumors to RT has long been the goal of PET imaging. The ability to differentiate responding from non-responding tumors before response is detectable through measurement of gross disease holds great promise. A question of fundamental importance is whether response should be assessed using the total tumor volume or the most aggressive regions of the initial tumor. Once a decision has been made as to the appropriate volume to assess response to RT, the question of how to align this volume on the subsequent imaging studies must be answered. To address this question, four arms were tested in specific aim 1.

The first two arms this specific aim attempted illuminate which volume delineation method should be preferred for predicting and assessing response to

RT with ^{18}F -FDG PET/CT. Total tumor volumes as well as the most metabolically active volumes defined on both PET and CT images were utilized to answer this question. Second, each arm was split into two arms to test the preferred alignment technique. A deformable image registration algorithm and a bony alignment method were utilized to align the pre-RT volumes across the serial imaging studies. Early identification of patients not responding to treatment may allow for more timely alteration in patient management.

Methods and Materials:

Patient Cohort

Patients were enrolled into an IRB approved protocol (LAB07-0043) at The University of Texas MD Anderson Cancer Center (MDACC) designed to assess and predict, prospectively, radiotherapy (RT) outcomes using serial ^{18}F -FDG PET/CT. Eligible adults had biopsy-proven stage III-IVb (American Joint Committee on Cancer) squamous cell carcinoma of the oropharynx, hypopharynx, or larynx with scheduled intensity modulated RT (IMRT) or chemoradiotherapy. Between 2005 and 2007, 107 patients were screened for enrollment.

All patients were scheduled for one pre-RT ^{18}F -FDG PET/CT study collected within four weeks prior to the beginning of treatment. Follow-up PET/CT studies were performed approximately 8 – 9 weeks after the completion

of treatment. All functional imaging studies were performed with integrated PET/CT systems.

Initial follow-up was performed eight to nine weeks after the completion of treatment. Follow-up radiographic, clinical, and pathologic evidence served as the gold standard for assessing patient outcome. Patients with no evidence of disease clinically or radiographically were continually monitored and considered as RT responders. Patients with residual or recurrent disease at the time of follow-up were considered RT non-responders.

RT and Systemic Treatment

Patients screened for inclusion followed two general therapeutic pathways. In one arm, patients received only IMRT; in the second arm, patients received concurrent systemic therapy with IMRT. Treatment plans were created with the Pinnacle³ treatment planning system (Philips Medical Systems, Andover, MA). The typical prescription dose was 70 Gy given 5 days a week in 2.12 Gy fractions. Other dose prescriptions used for this cohort were 1.8 Gy/fx, 2.0 Gy/fx, and 2.2 Gy/fx. Patients were not stratified based upon varying doses per fraction.

Functional Imaging

PET imaging studies were performed with the integrated GE Discovery (GE Medical Systems, Milwaukee, Wisconsin) PET/CT systems. Following collection all images were transferred to the institutional digital archival system.

Patient preparation for F-FDG PET/CT imaging was crucial to establish uniform collection conditions of all functional studies. Patients were instructed to fast for at least four hours prior to injection of F-FDG radiotracer. It is important for patients to maintain a high protein-low carbohydrate diet, abstaining from sugary drinks, but still maintaining hydration. Reducing the blood glucose concentration will reduce uptake competition between injected FDG and consumed carbohydrates. Prior to imaging, serum glucose levels were measured and patients found to have levels greater than 200 mg/dL were not imaged. Patient height (m) and weight (kg) were also recorded. Imaging subjects were then instructed to lie supine and minimize movement to reduce the accumulation of FDG in muscle.

Tracer was injected intravenously with a 20 – 23 gauge intracatheter. Institutional guidelines specify the injection of 5 – 20 mCi of ^{18}F -FDG followed by a saline flush of 10 – 20 cm³. Following injection, procedural recommendations specify a 30 – 45 minute waiting period to allow the tracer to be absorbed throughout the patient's tissues. American College of Radiology Imaging Network (ACRIN) guidelines stipulate that imaging should begin within 60 ± 10 minutes. Additionally, all follow-up PET/CT studies should match the injection to imaging time as closely as possible, varying no more than 10 minutes.

Anatomic Imaging

In addition to PET/CT imaging studies, patients also had contrast enhanced CT studies performed. The first was collected prior to the start of

treatment and the follow-up was collected 8 weeks after the completion of treatment. All images were reviewed by a board certified radiologist and classified as representing complete, partial, stable, or progressive response [5]. The interpretation of contrast enhanced CT images was performed without knowledge of the PET/CT results.

Post-Treatment Surveillance

Post-RT surveillance followed a well defined pathway. Initial follow-up was scheduled for 8 weeks after the completion of treatment. Subsequently, patients were assessed every 3 – 4 months for the first two years with clinical examination and contrast enhanced CT studies of the head and neck. Patients with no evidence of disease were monitored with regularly scheduled visits.

Patients with unequivocal incomplete nodal response had unilateral or bilateral neck dissection performed. Equivocal findings were assessed by ultrasound-guided fine needle aspiration. Patients with negative findings were followed with regularly scheduled visits. Additionally, patients with positive findings on clinical or radiographic examination for recurrent primary tumors had salvage surgical resection performed.

Pathologic Tissue Assessment

A single board certified pathologist performed all step-sectioning and evaluation of neck dissection samples. This methodology is detailed in a previous report by Moeller et al [34]. Briefly, nodes measuring greater than 1 cm

were bisected along the longest diameter. Both halves were then sectioned with additional levels cut at 150 μm intervals. If a node measured less than 1 cm, one section was obtained with additional levels at 100 μm intervals. At least one subcapsular squamous cell cluster was required for positive metastatic involvement.

Tumor Volume Delineation

The ability to differentiate responding and non-responding primary tumors was investigated for SUV derived in two general methods. First, the frequently utilized SUV_{Max} , the maximum pixel value for the most FDG-avid regions, was collected from user defined regions that fully encompassed the most FDG-avid tumor regions on pre-RT PET/CT images. Regions similar in size and anatomic relation were drawn on post-RT images for post-therapeutic assessment.

Next, SUV_{Mean} was derived from three volumetric techniques. First, the gross tumor volume (GTV) from the clinically delivered treatment plan was used as the ROI. The GTV volume in head and neck treatments at our institution contain both primary tumor and nodes. However, for the purposes of this analysis primary tumor volumes were separated from involved nodes. GTV contours created on planning CT images were used to collect baseline and follow-up SUV_{Mean} . Although not created on PET images, the GTV region-of-interest (ROI) takes advantage of the inherent registration between PET and CT images collected with integrated scanner systems. SUV derived from this metric represented a CT defined total tumor volume.

Next, a volume defined exclusively on the pre-RT PET images was utilized. The volume was defined with an auto-contouring algorithm created for use with the Pinnacle³ treatment planning system. The source code is given in the appendix under *PET50Threshold.script*. After the PET images were imported into the Pinnacle³ treatment planning system and assigned an appropriate primary, CT, image set, the algorithm required the user to define a region that fully resolved the most FDG-avid region of the image that is of interest. This contour is denoted as “FindMaxPixel,” in the source code. Most importantly, prior to running the script, the destination contour had the primary volume set as the pre-RT PET image. After running the script, an isocontour was created with the name “50Primary.” This contour encompassed all pixels in the FDG-avid region that are 50% of the maximum pixel value or greater. This value was selected because of its frequent use in the literature [36, 78, 79]. The contour was then transferred to the corresponding volume on the primary image set, in this case the pre-RT CT. This volume was then deformably mapped, as describe above, to post-RT PET/CT images where it was used to collect the average SUV. SUV derived from this metric represented a PET defined total tumor volume.

Finally, SUV_{Peak} for the FDG-avid regions at the site of the primary tumor were determined. SUV_{Peak} is defined as an approximately 1 cm³ volume centered on the maximum pixel of an FDG-avid region with a diameter of 1.2 cm. Utilizing the script *SUVpeak.m*, SUV_{peak} volumes were defined on pre-RT PET/CT images. The approximately 1 cm³ volume was then deformably

mapped to post-RT PET/CT images in a procedure analogous to that described above. SUV derived from this metric represented the most metabolically active volume.

SUV values of responding and non-responding patients derived from each technique were then compared. Nonparametric Mann-Whitney rank sum tests were used to infer whether the SUV values of responding and non-responding patients were different. Pre- and post-RT absolute SUV metrics as well as relative change in SUV defined as,

$$\text{Fractional SUV} = \frac{\text{Post - RT SUV}}{\text{Pre - RT SUV}}$$

were tested. Statistically significant differences are considered to be those with $P < 0.05$.

Image Registration

The alignment of tumor volumes was accomplished with two techniques. First, pre-RT tumor volumes were mapped to serial imaging studies with an enhanced “demons” algorithm. Second, pre-RT volumes were registered to serial imaging studies with rigid alignment.

The first alignment method employed a deformable image registration technique [80, 81]. Based upon the “demons” algorithm originally proposed by Thirion et al., it is based upon CT image intensity (Hounsfield units) [82].

Intensity differences between a static and moving image are minimized iteratively producing a displacement vector for each voxel. The displacement vector field is then applied to the reference image, effectively deforming or mapping voxels from the reference to the target image. Implementation of a multi-resolution approach and the addition of an active force derived from the intensity gradient information in the target image, have allowed the enhanced “demons” algorithm to greatly improve the registration quality over the original implementation. Wang et al. have shown in mathematical and phantom studies the overall registration error of the enhanced “demons” algorithm to be approximately 1 mm [81].

For the alternate registration technique, a CT-to-CT bony alignment algorithm was utilized [83, 84]. By exploiting an ROI defined on the planning CT images, the algorithm matches the same image feature in a target image set by minimizing a cost function. The cost function used is the mean absolute difference in CT numbers, computed on a voxel-by-voxel basis, between the planning and, in this case, either the pre- or post-RT PET/CT. The ROI used for bony alignment was the C2 vertebra. The C2 vertebra was selected because of its proximity to the multiple disease sites for the current cohort. Additionally, in one study by Zhang et al., C2 was found to have the smallest systematic and random daily setup variations in the anterior-posterior (AP) and superior-inferior (SI) directions, making it a stable alignment point [85].

Volume Alignment Workflow

The following methodology details the steps required to utilize the CT Assisted Targeting for Radiotherapy (CAT) software to collect SUV for a volume of interest mapped from a reference image set to a daily image set. Briefly, the methodology involved first contouring the C2 vertebrae for each patient for initial rigid alignment. Second, tumor volumes created within the Pinnacle Treatment planning system and MATLAB were deformed, separately, to a daily CT image set. For this particular project, the reference image set was the planning CT (PCT) and the daily image sets were the pre- and post-RT PET/CT image sets. After all contours were deformed, the SUV and dose were collected from each volume of interest. The output format was a text file with an (x,y,z) voxel location within the image volume and a numeric value for the voxel (SUV or dose). The workflow will now be explained in detail and will be of particular interest to future investigators using the CAT software for PET/CT research.

Prior to any post-processing, all serial image sets were collected. This was accomplished using the ClinicStation software platform. Within this platform, all serial studies were reviewed, selected, and then exported as DICOM images to the internet database Evercore. After the images were downloaded from Evercore, folders containing the DICOM images were renamed to CT01, CT02, PET01, or PET02 using the *scanorder.m* script to more easily identify the image study chronology. While created for this specific task

the script *scanorder.m* may be of use to other investigators using multiple image sets. This script is given in the appendix.

Next, all patient treatment plans were collected. Within Philips' Pinnacle³ Treatment Planning System (version 7.6c), the axis (C2 vertebrae) was contoured for every patient. The axis was contoured for each patient for initial three dimensional bone alignments. Volumes of interest obtained from the original treatment plan included the GTV and PET threshold algorithm. Although not created on the planning CT images, the PET based threshold contour was created using the Pinnacle³ treatment planning system.

Following the collection of all imaging studies and volumes of interest, data were imported into the CAT software workspace. This includes treatment plans and all functional imaging studies. Within the CAT workspace, the reference image (PCT) was first rigidly aligned to the daily image set (pre- or post-RT PET/CT). In the case of the PET-based threshold volume the reference image set was actually the pre-RT PET/CT image set, but the workflow was identical. The initial bony alignment utilized the algorithm specified in the previous section. After a global minimum for the cost function was achieved, translation shifts in the left/right, anterior/posterior, and superior/inferior were applied to the volumes of interest thereby aligning them to the daily image set. For the bony alignment arm of this aim, no further registration steps were required and SUV was collected for pre- and post-RT PET/CT images.

For the deformable registration arm, additional alignment steps were required. Next, a user-defined deformation volume was selected to completely resolve the site of primary tumor with boundaries stopping approximately inferiorly at the level of the suprasternal notch and continuing to the most superior slice of the PET image set. The enhanced demons algorithm was then utilized to map all voxels within the user defined deformation volume from the reference image set to the daily image set. This resulted in a vector displacement field that mapped the reference image (planning CT) separately to the daily images (pre- and post-RT PET/CT) images. Visual inspection was utilized to compare the daily image set with the resulting deformed reference image set. The deformable transformations were then applied to the volumes of interest in order to map the regions to either the pre- or post-RT PET/CT images. Finally, SUV was collected from the serial PET/CT images using the mapped volumes of interest.

Calculating SUV in the CAT Workspace

Tumor glucose use in this work was quantified with SUV normalized to total body mass [86]. In order to calculate SUV, the injected dose activity in MBq was extracted from the DICOM header and decay corrected to the time of imaging by the following relation,

$$\text{Injected Dose}_{\text{Imaging time}} = \text{Injected Dose} * 0.5^{\frac{\text{Injection to imaging time}}{6588}}$$

where the injected doses are in units of MBq and the denominator in the exponent is the half life of ^{18}F . Next, each pixel in the image was multiplied by the rescale slope in the DICOM header to account for the restricted maximum pixel value of a GE PET image, 32, 7636 Bq/mL. Finally, the patient mass was extracted and SUV was computed according to the following relation,

$$SUV = \frac{\text{Pixel value [Bq/mL]} * \text{Rescale slope} * \text{Patient mass [g]}}{\text{InjectedDose}_{\text{Imagingtime}}}$$

Additionally, the CAT software platform capability to calculate SUV was expanded to include lean body mass and body surface area normalization [87].

Receiver Operating Characteristic Analysis

Receiver operating characteristic (ROC) analyses were performed to compare the ability of each metric described above to predict response to RT. Classification of patients as normal or abnormal was based upon response to RT as determined through pathology and continued clinical surveillance. Median follow-up time was approximately two years post-RT. Software from SPSS an IBM Company (Chicago, IL) was utilized to perform ROC analyses. The ROC curve is a plot of the true positive fraction, sensitivity, versus the false positive fraction, (1 – specificity). In other words, the sensitivity is the fraction of

cases that a clinician calls abnormal that are actually abnormal as determined by an alternate gold standard. The specificity is the fraction of cases that a clinician calls normal that are actually normal. After ROC curves were created for all SUV metrics utilized, decision thresholds were established based upon tabulated data points of each ROC curve. The point along the ROC curve that maximized sensitivity and minimized $(1 - \text{specificity})$ was selected as the decision threshold for a particular metric. In addition to calculating the true positive fraction and true negative fraction for each technique (sensitivity and specificity), positive and negative predictive values were computed (PPV and NPN). The PPV refers to the probability that a patient is actually abnormal or a non-responder when classified as such. On the other hand, NPV relates to the probability that the patient is actually normal or responded to RT when classified accordingly.

Finally, the ability of each metric to discriminate responding and non-responding patients was compared. The ROC curve is constructed by determining the sensitivity and specificity at various decision threshold values. Therefore, the area under the ROC curve (AUC) will depend upon how well the metric is able to discriminate responding and non-responding patients and help to quantify the usefulness of the metric. In the case of pure guessing, $\text{AUC} = 0.5$. If AUC is less than 0.5, the metric performs worse than guessing. Software from MedCalc (Mariakerke, Belgium) was used to perform non-parametric tests to compare the AUC for each metric [88]. Statistically significant differences are considered to be those with $P < 0.05$.

Results:

Patient Characteristics

Following the enrollment period lasting from November 2005 through May 2007, 107 patients were screened for enrollment. 15 patients were rejected for multiple reasons. These included: prior surgery at the site of primary tumor (n=10), imaged off protocol (n=3), and withdrawal of consent (n=2). This left a total of 92 patients. Of these 22 were referred for surgical evaluation at the time of first follow-up. Of those, 16 had no evidence of disease, responders, and 6 were found to have residual disease, non-responders. For the clinical follow-up arm, there were 6 patients with evidence of residual disease, non-responders, and 64 patients that responded to treatment. Because this analysis also made use of the archived radiation treatment plan, various archiving errors further reduced the useable population of 92, n=14. Additionally, patients with extreme anterior or posterior rotation of the head and neck were excluded for the analysis because of the difficulty in registering such patients, n = 10. This left 49 responding patients and 6 non-responding.

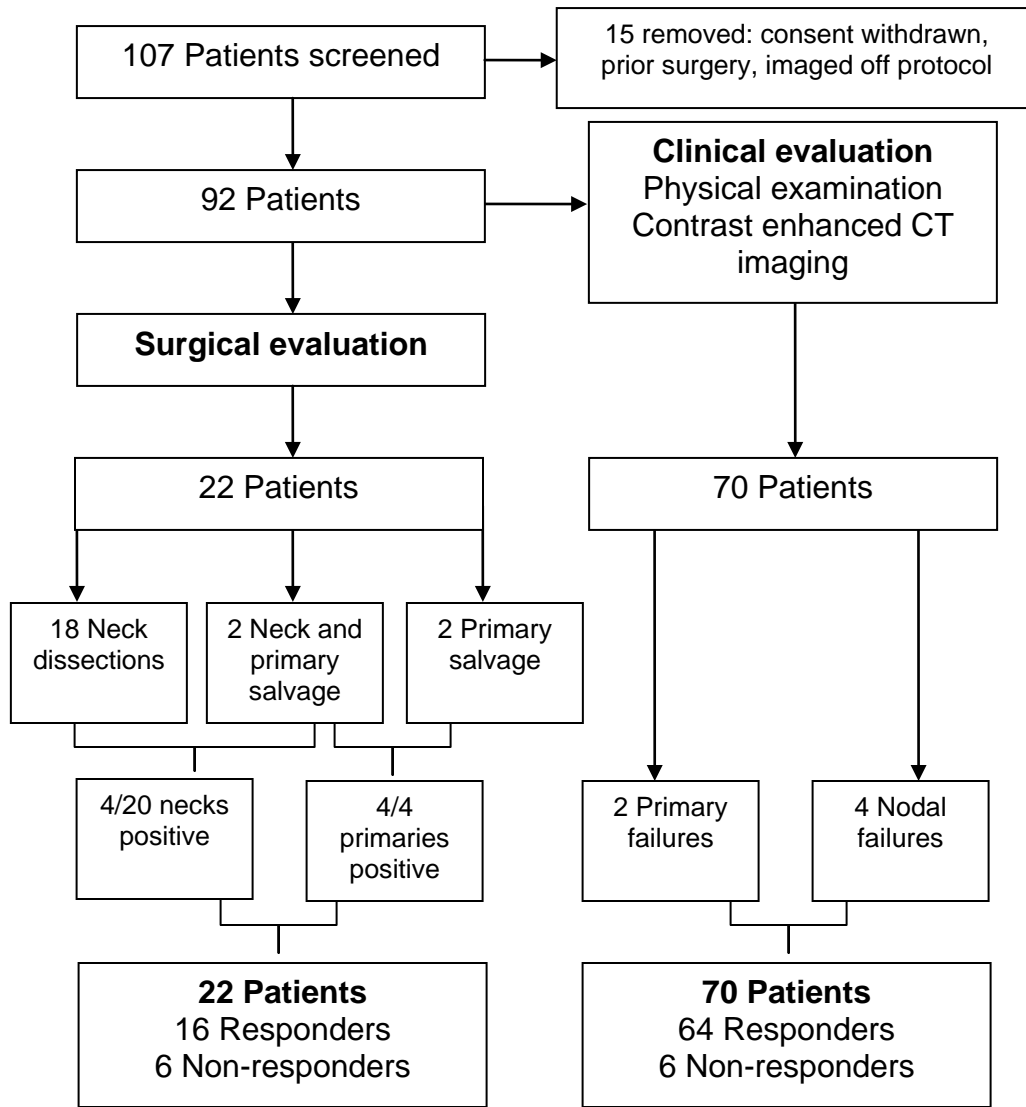
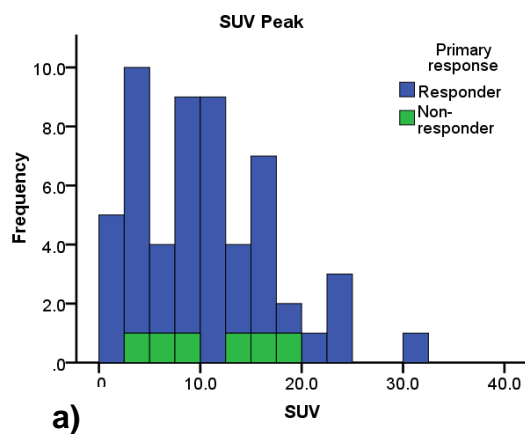


Figure 3.1 Post Treatment Surveillance Workflow

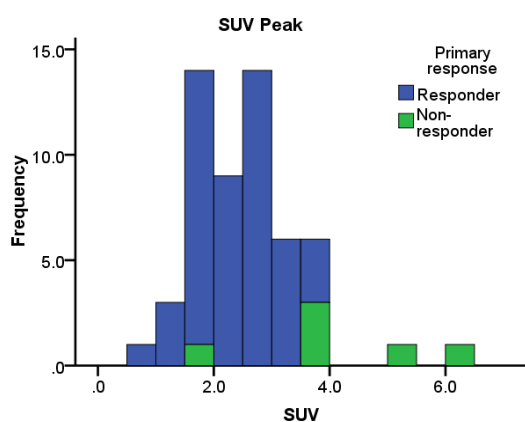
SUV and Treatment Response

SUV was first analyzed to determine if any associations with response to treatment existed. Pre- and post-RT absolute SUV metrics were tested as well as relative change in SUV. Summaries of all results are presented in table 3.1.

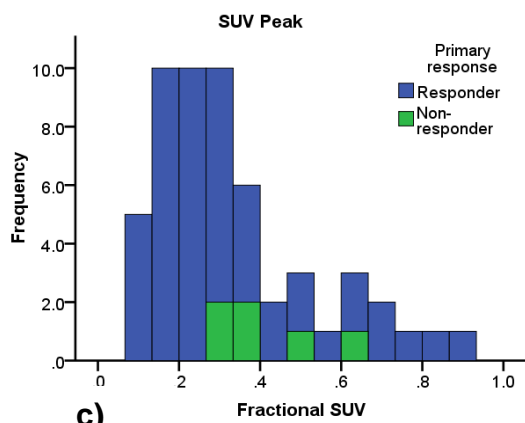
The SUV_{Peak} contour was defined on pre-RT PET images as an approximately 1 cm^3 volume. This volume was then deformably mapped to the post-RT PET/CT study. The calculated pre-RT SUV_{Peak} of the most metabolically active volume of responding primary tumors was found to be 10 ± 6.7 . In non-responding patients, the pre-RT SUV_{Peak} was 13 ± 5.0 . No statistically significant differences were found between the SUV values in these two groups ($P > 0.05$). The distribution of SUV_{Peak} values are shown in figure 3.2a. However, statistically detectable differences were found between the post-RT SUV_{Peak} values of responding tumors (2.4 ± 0.7) and non-responding (4.6 ± 1.3) patients ($P < 0.01$). The histogram distribution for this metric is shown in figure 3.2b. The separation between SUV_{Peak} values of responding and non-responding patients is clearly visible. Additionally, the relative change between the pre- and post-RT SUV_{Peak} , expressed as fractional SUV was evaluated for this ROI. No statistically discernable differences were found between fractional SUV_{Peak} values for responding (0.3 ± 0.2) and non-responding (0.4 ± 0.1) patients ($P > 0.05$). The distributions of these values are illustrated in figure 3.2c.



a)



b)

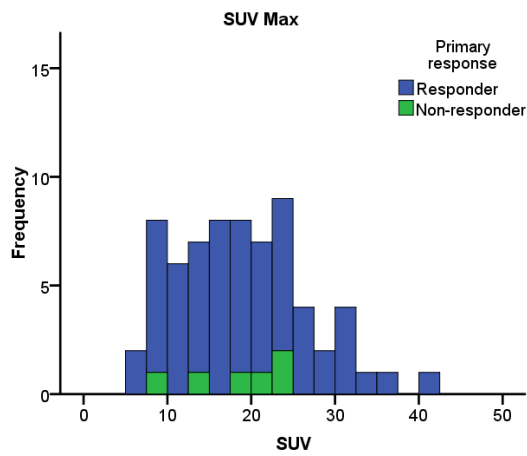


c)

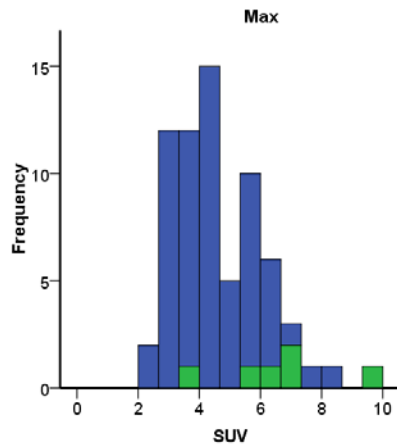
Figure 3.2 Deformably aligned SUV_{Peak}

Histogram distributions of (a) pre- (b) post-RT , and (c) fractional SUV_{Peak} derived from the 1 cm³ SUV_{Peak} contour created on pre-RT PET images and deformed to post-RT PET images.

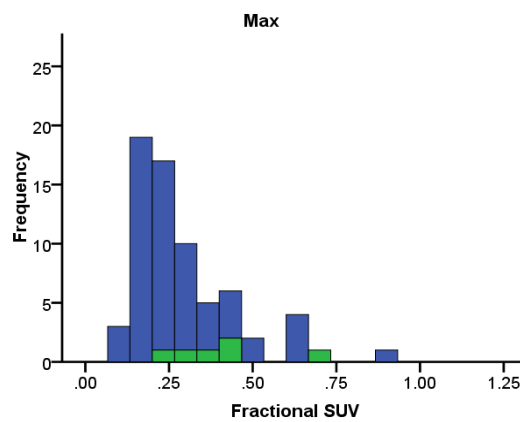
The SUV_{Max} metric was evaluated next. SUV_{Max} , a single pixel value, is hypothesized to represent the maximum tumor burden. This value was determined from a user defined volume that fully circumscribed the FDG-avid regions. The calculated pre-RT SUV_{Max} was found to be 19 ± 7.7 in patients who responded to RT. In non-responding patients, the pre-RT SUV_{Max} was 18 ± 5.9 . No statistically significant differences were found between the SUV values in these two groups ($P > 0.05$). The distribution of pre-RT SUV_{Max} values are shown in figure 3.3a. However, statistically detectable differences were found between the post-RT SUV_{Max} values of responding tumors (4.5 ± 1.3) and non-responding (6.7 ± 2.0) patients ($P < 0.01$). The histogram distribution for this metric is shown in figure 3.3b. The separation between SUV_{Max} values of responding and non-responding patients is clearly visible, yet not as pronounced as seen for the SUV_{Peak} contour. The relative change between the pre- and post-RT SUV_{Max} , was found to be statistically significant ($P < 0.05$). The distribution of fractional SUV_{Peak} for responding (0.3 ± 0.2) and non-responding (0.4 ± 0.2) patients is shown in figure 3.3c.



a)



b)



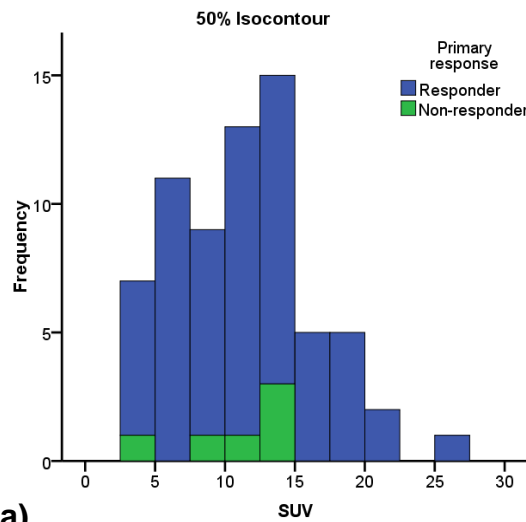
c)

Figure 3.3 SUV_{Max}

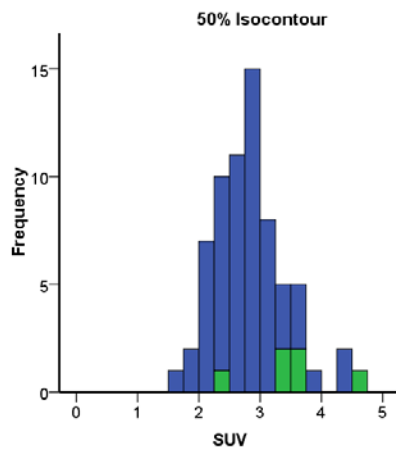
histograms

Histogram distributions of (a) pre- (b) post-RT, and (c) fractional SUV_{Max} derived from a user defined region that fully resolved the FDG avid region.

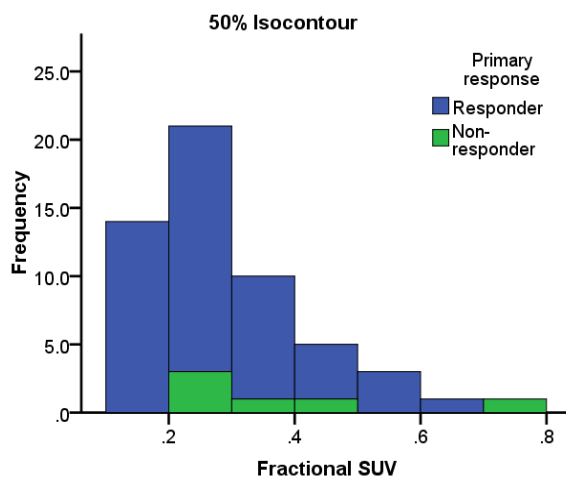
The 50% isocontour volume was defined on pre-RT PET images and encompassed all pixels in the FDG-avid volume that were within 50% of the maximum pixel value. This volume was then mapped to the post-RT PET/CT study. The calculated pre-RT SUV for the threshold PET-based contour was found to be 11 ± 5.1 . In non-responding patients, SUVMean for the 50% isocontour was 11 ± 5.9 . No statistically significant differences were found between the SUVMean values in these two groups ($P > 0.05$). The distribution of SUVMean values for the threshold ROI are shown in figure 3.4a. However, statistically detectable differences were found between the post-RT SUVMean values of responding tumors (2.8 ± 0.5) and non-responding (3.4 ± 0.7) patients ($P < 0.05$). The histogram distribution for this metric is shown in figure 3.4b. Finally, the relative change between the pre- and post-RT SUVMean for this ROI, expressed as fractional SUV was evaluated. No statistically discernable differences were found between responding fractional SUVMean (0.3 ± 0.2) and non-responding (0.4 ± 0.2) fractional SUVMean at the 95% confidence level. The distributions of these values are illustrated in figure 3.4c.



a)



b)

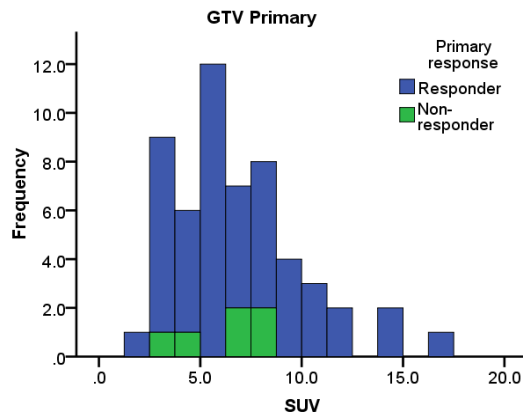


c)

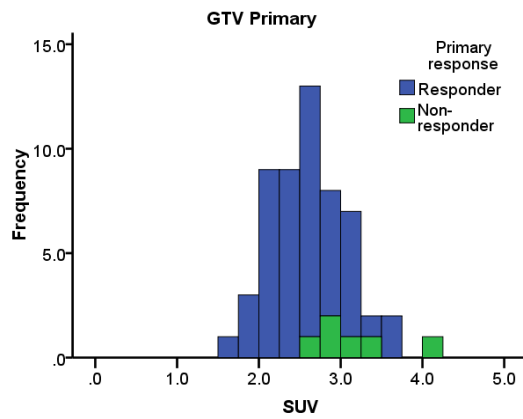
**Figure 3.4 Deformably
aligned PET-based threshold
contour**

Histogram distributions of (a) pre- (b) post-RT , and (c) fractional SUV_{Mean} derived from deformably mapped threshold contour that encompasses all pixels within 50% of SUV_{Max} .

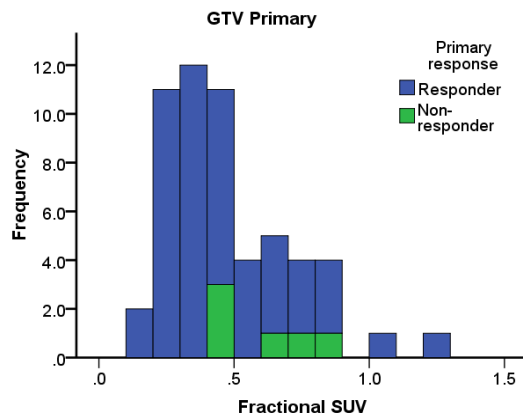
The final contour to be evaluated was the CT-based GTV. The GTV contours created for this study were created on planning CT images and then mapped separately to pre- and post-RT PET/CT images, where they were utilized as ROIs to determine SUV. The calculated pre-RT SUV for the gross tumor volume of patients who ultimately responded to treatment was 6.7 ± 3.3 . In non-responding patients, the pre-RT SUV was 5.8 ± 2.1 . No statistically significant differences were found between the SUV values in these two groups ($P > 0.05$). The distribution of SUVMean values derived for the total tumor volume is shown in figure 3.5a. However, statistically detectable differences were found between the post-RT SUV values of responding tumors (2.6 ± 0.4) and non-responding (3.2 ± 0.7) patients ($P < 0.05$). The histogram distribution for this metric is shown in figure 3.5b. The separation between SUVMean values of responding and non-responding patients is clearly visible, but marked overlap between the two groups is seen. Additionally, the relative change between the pre- and post-RT SUV, expressed as fractional SUV was evaluated for this ROI. No statistically discernable differences were found between responding fractional SUVPeak (0.5 ± 0.3) and non-responding (0.6 ± 0.2) at the 95% confidence level. The distributions of these values are illustrated in figure 3.5c.



a)



b)



c)

Figure 3.5 Deformably aligned GTV contour

Histogram distributions of (a) pre- (b) post-RT , and (c) fractional SUV_{Mean} derived from the GTV contour created on planning CT images and deformably mapped to pre- and post-RT PET/CT images.

	Responder			Non-responder		
	Pre-RT	Post-RT	Fractional SUV	Pre-RT	Post-RT	Fractional SUV
SUV_{Max}	19 ± 7.7	4.5 ± 1.3	0.3 ± 0.2	18 ± 5.9	6.7 ± 2.0	0.4 ± 0.2
SUV_{Peak}	10 ± 6.7	2.4 ± 0.7	0.3 ± 0.2	13 ± 5.0	4.6 ± 1.3	0.4 ± 0.1
50% Isocontour	11 ± 5.1	2.8 ± 0.5	0.3 ± 0.2	11 ± 3.9	3.4 ± 0.7	0.4 ± 0.2
GTV primary	6.7 ± 3.3	2.6 ± 0.4	0.5 ± 0.3	5.8 ± 2.1	3.2 ± 0.7	0.6 ± 0.2
Background	4.9 ± 1.0	4.6 ± 1.0	0.9 ± 0.2	4.3 ± 1.1	4.0 ± 1.0	0.9 ± 0.3

Table 3.1 Average SUV values derived from deformably mapped volumes for responding and non-responding patients.

Tabulated here are absolute and fractional SUV metrics for the cohort stratified by primary tumor response. One standard deviation is presented along with each mean value. Statistically discernable differences between responding and non-responding patients are shown in bold face type (P<0.05).

Predicting RT Outcomes with SUV

Absolute post-RT SUV was identified as the superior metric to differentiate patients responding to RT from those who did not. Next, post-RT SUV decision thresholds were determined for each technique based upon tabulated data from each ROC curve. In figure 3.6, the ROC curves for post-RT SUV from each technique are shown. For reference, the curve representing performance of pure guessing is plotted as a dashed line. Additionally, the sensitivity and (1 – Specificity) for the post-RT contrast enhanced CT is shown as a single data point for comparison. The binary end point, response to RT, was defined as a reading of complete radiographic response. All other radiographic endpoints: partial response, stable disease, and progressive disease were defined as non-responders. Post-RT contrast enhanced CT was found to have a sensitivity of 80% and a specificity of 92%. Radiographic analysis yielded one false positive. Overall, CT had a positive predicted value of 50%.

Decision thresholds that maximize sensitivity and minimize (1 – specificity) for each technique were selected from tabulated ROC curves. For the SUV_{Max} metric, a post-RT absolute SUV decision threshold value of 6.4 was selected. Based upon this cutoff, sensitivity (80%) and specificity (94%) were calculated. This cutoff point led to one false negative. The positive predictive value of post-RT SUV_{Max} was calculated to be 57%. This result is slightly better

than what was found for contrast enhanced CT and notable for performing better than guessing.

The best performing technique was the SUV_{Peak} metric. A post-RT SUV decision threshold of 3.5 was selected from the tabulated ROC analysis. This cutoff point yielded the highest sensitivity of any method, 100%, and a very high specificity, 94%. Notably, this was the only technique to have zero false positives. The positive predictive value was 63% and the negative predictive value was 100%, the highest values for all techniques investigated.

The 50% threshold and GTV contour both performed poorly. For the PET-based threshold contour, a decision threshold of 3.3 was selected. Utilizing this cutoff, the sensitivity was found to be no better than contrast enhanced CT, 80%, and the specificity was found to be worse, 88%. Similar to CT evaluation, this analysis resulted in one false negative. While the negative predictive value was high, 98%, the positive predictive value was less than 50%. Results for SUV_{Mean} derived from the mapped GTV contour were similar. A decision threshold of 2.8 was identified. This resulted in one false negative. The sensitivity was identical to all other techniques, aside from SUV_{Peak} , but ultimately proved to have the worst specificity, 69%. Furthermore, the positive predictive value of this technique was the lowest at 21%. Table 3.2 displays the results of all ROC analyses.

The SUV_{Peak} method resulted in the largest area under the ROC curve, 0.98 (0.89 – 0.99, 95% C.I.). Both SUV_{Max} and SUV_{Mean} derived from the 50%

threshold contour had areas of 0.93 (0.82 – 0.98 95% C.I.). Finally, SUV_{Mean} determined for the GTV volume had the smallest area, 0.8 (0.67 – 0.90 95% C.I.). First, consider the SUV_{Peak} metric that resulted in the highest positive and negative predictive values. Statistically significant differences in the area under the ROC curve were found between the SUV_{Peak} metric and SUV_{Mean} derived from the GTV contour ($P < 0.05$). No statistically significant differences were found between the area under the SUV_{Peak} curve and the area under the curves of all other metrics ($P > 0.05$). Pair-wise comparisons between all other metrics found no significant differences in the area under each curve ($P > 0.05$).

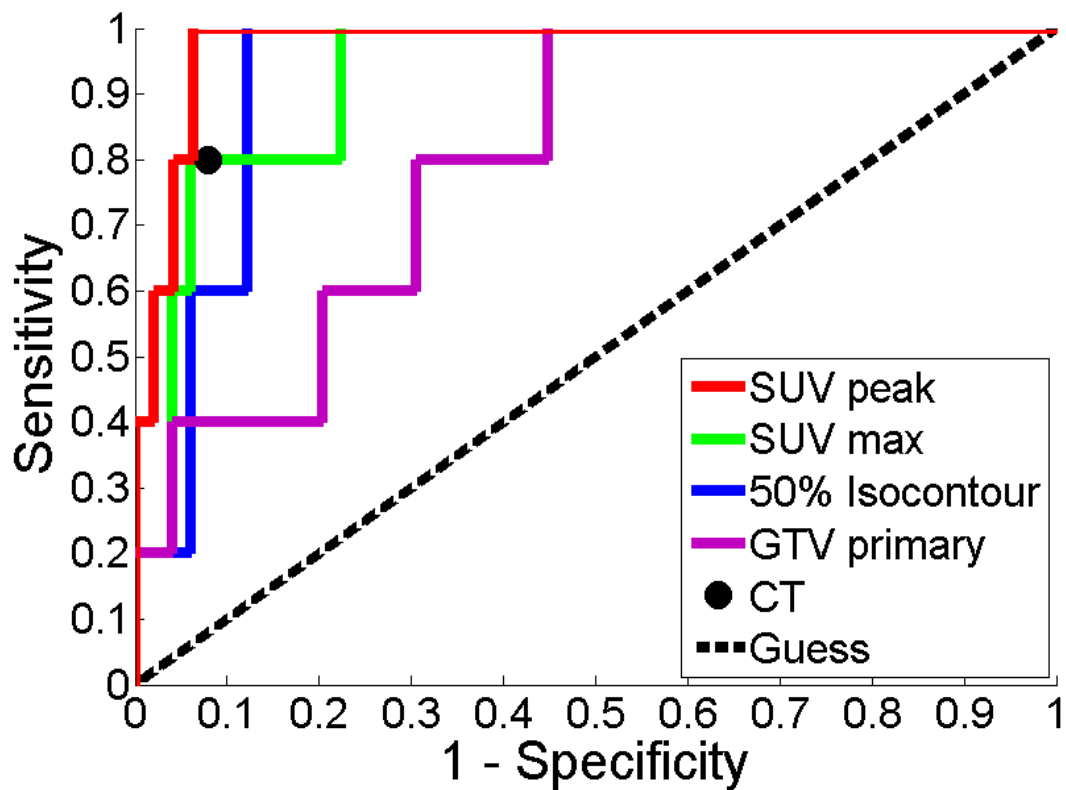


Figure 3.6 ROC curves for the deformably aligned volumes

ROC curves for SUV_{Max} and SUV_{Mean} derived from the GTV contoured on planning CT images and a PET-based 50% threshold contour and SUV_{Peak} are shown. The single data point represents the sensitivity and specificity of contrast-enhanced CT for the study.

	CT	SUV _{Max}	SUV _{Peak}	50% Isocontour	GTV Primary
Decision threshold		6.4	3.5	3.3	2.8
True positives	4	4	5	4	4
False negatives	1	1	0	1	1
True negatives	45	46	46	43	34
False positives	4	3	3	6	15
Sensitivity	0.80	0.80	1.00	0.80	0.80
Specificity	0.92	0.94	0.94	0.88	0.69
Positive predictive value	0.50	0.57	0.63	0.40	0.21
Negative predictive value	0.98	0.98	1.00	0.98	0.97

Table 3.2 Accuracy of response prediction with deformably mapped volumes

Cutoff values were selected from ROC curves to classify patients as normal or abnormal. The raw numbers for each error type are tabulated along with calculated sensitivity, specificity, positive predictive values, and negative predictive values. The superior metric is denoted in bold face type.

Correlating SUV Derived from Bony Alignment Contours to Treatment Response

While the first major arm of this aim investigated the ability of SUV derived from mapped contours to assess and predict response to RT, the second arm investigated the impact of improved registration on treatment response monitoring. For the second major arm of this specific aim, SUV was derived from pre-treatment volumes aligned to the C2 vertebra on pre- and post-RT PET/CT images. Pre- and post-RT absolute SUV metrics, as well as relative change in SUV, derived from the rigidly aligned contours were analyzed for correlations to response to RT. Summaries of all results are presented in table 3.3.

Post-RT absolute SUV_{Peak} derived from contours mapped to pre- and post-RT PET/CT studies was the best performing metric in the previous analysis. The calculated pre-RT SUV of the most metabolically active volume of responding primary tumors was found to be 10 ± 6.7 . In non-responding patients, the pre-RT SUV_{Peak} was 13 ± 5.0 . These values are identical to those from the previous section as this contour is defined on pre-RT PET/CT images. Therefore, no statistically significant differences were found between the SUV values in these two groups ($P > 0.05$) as discussed above. Contrary to the previous SUV_{Peak} results, statistically detectable differences were not found between the post-RT SUV_{Peak} values of responding tumors (2.6 ± 0.6) and non-responding (2.8 ± 0.8) patients ($P > 0.05$). The histogram distribution for this

metric is shown in figure 3.7b. For comparison, the histogram distribution for the deformably mapped SUV_{Peak} volume is shown in figure 3.7a. It is clear from the distributions that bony alignment of the pre-RT SUV_{Peak} volume was not sufficient to capture the FDG-avidity of the residual tumor. Additionally, the relative change between the pre- and post-RT SUV_{Peak} , expressed as fractional SUV was evaluated for this ROI. Once again no statistically discernable differences were found between responding fractional SUV_{Peak} (0.4 ± 0.4) and non-responding (0.3 ± 0.1) ($P>0.05$).

The PET-based threshold contour was analyzed next. The calculated pre-RT SUV for the threshold PET-based contour was found to be 11 ± 5.1 . In non-responding patients, SUV_{Mean} for the 50% isocontour was 11 ± 3.9 . As mentioned above, these volumes are defined on the pre-RT PET/CT images and the results of the previous analysis will not alter in the present scenario. In contrast to the analyses with deformed contours, statistically detectable differences were not found between the post-RT SUV_{Mean} values of responding tumors (2.0 ± 0.4) and non-responding (2.1 ± 0.3) patients ($P>0.05$). The histogram distribution for this metric is shown in figure 3.8b. In figure 3.8a, the distribution of SUV_{Mean} derived from the mapped threshold contour is shown for reference. Of note is the shift toward lower SUV values for the rigidly aligned contours, indicating incomplete resolution of the FDG-avid volume. Finally, the relative change between the pre- and post-RT SUV_{Mean} for this ROI, expressed as fractional SUV was evaluated. No statistically discernable differences were

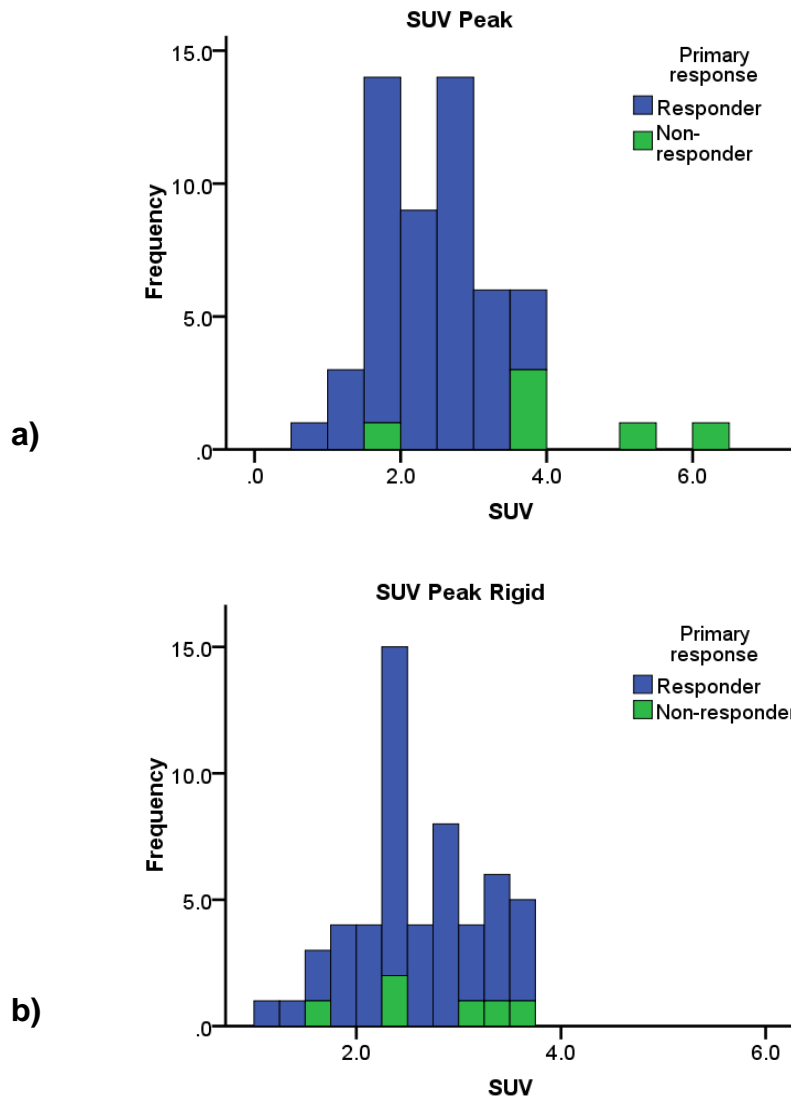


Figure 3.7 Bony and deformed SUV_{Peak} distributions

Histogram distributions for deformably (a) and rigidly (b) aligned PET-based SUV_{Peak} contour.

found between responding fractional SUV_{Mean} , 0.2 ± 0.1 , and non-responding, 0.2 ± 0.1 , fractional SUV_{Mean} , $P > 0.05$.

The final contour to be evaluated was the bony aligned CT-based GTV. The calculated pre-RT SUV for the gross tumor volume of patients who ultimately responded to treatment was 3.5 ± 1.3 . In non-responding patients, the pre-RT SUV was 5.6 ± 2.5 . No statistically significant differences were found between the SUV values in these two groups ($P > 0.05$). As seen for all rigidly aligned contours, no statistically detectable differences were found between the post-RT SUV values of responding tumors (1.9 ± 0.3) and non-responding (2.1 ± 0.3) patients ($P > 0.05$). The histogram distribution for this metric is shown in figure 3.9b. Additionally, the relative change between the pre- and post-RT SUV, expressed as fractional SUV was evaluated for this ROI. No statistically discernable differences were found between responding fractional SUV_{Peak} (0.6 ± 0.2) and non-responding (0.5 ± 0.3) ($P > 0.05$).

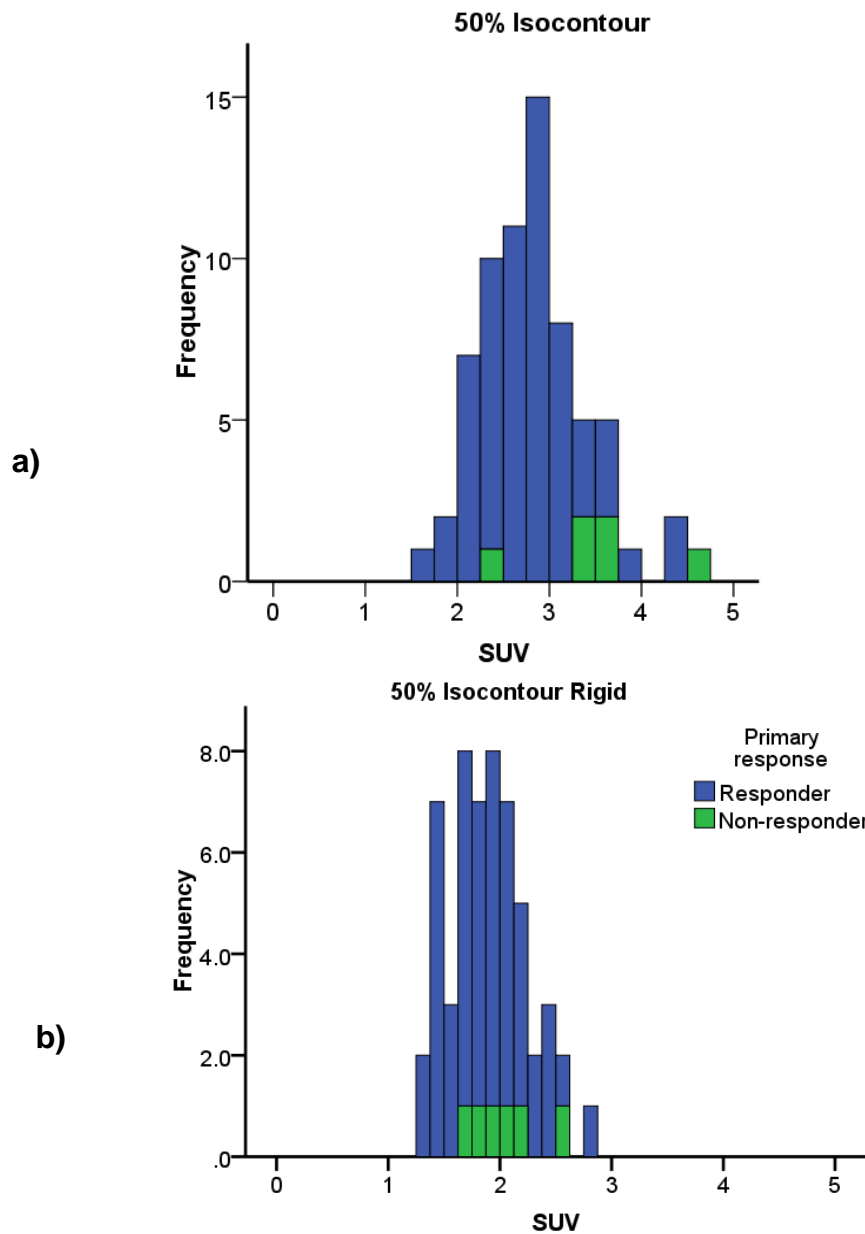


Figure 3.8 Bony and deformed PET threshold contour

Histogram distributions for deformably (a) and rigidly (b) aligned PET-based 50% threshold contour.

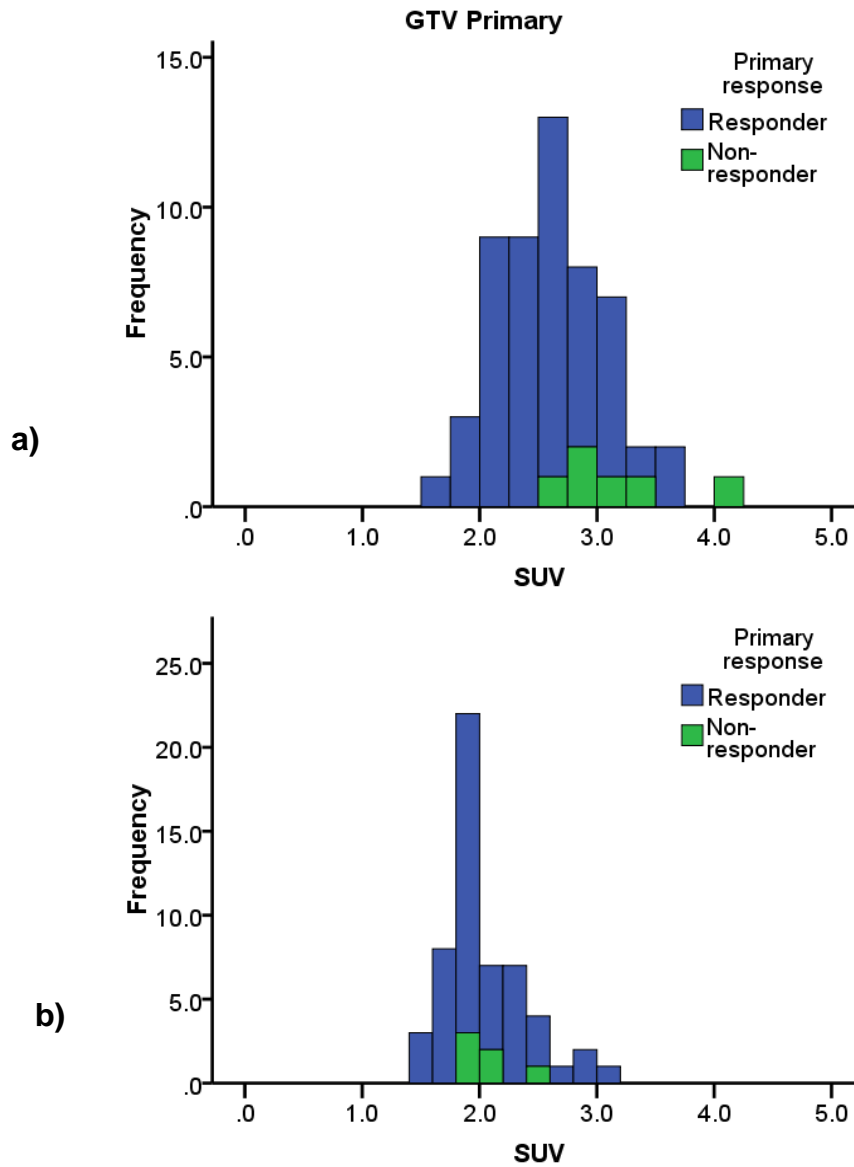


Figure 3.9 Bony and deformed GTV

Histogram distributions for deformably (a) and rigidly (b) aligned CT-based GTV contour.

	Responder			Non-responder		
	Pre-RT	Post-RT	Fractional SUV	Pre-RT	Post-RT	Fractional SUV
SUV _{Peak}	10 6.7	2.6 0.6	0.4 0.4	13 5.0	2.8 0.8	0.3 0.1
50% Isocontour	11 5.1	2.0 0.4	0.2 0.1	11 3.9	2.1 0.3	0.2 0.1
GTV primary	3.5 1.3	1.9 0.3	0.6 0.2	5.6 2.5	2.1 0.3	0.5 0.3

Table 3.3 Average SUV values for responding and non-responding patients from rigidly aligned volumes.

Tabulated here are absolute and fractional SUV metrics derived from rigidly aligned ROIs for the cohort stratified by primary tumor response. One standard deviation is presented along with each mean value. No statistically discernable differences were found between SUV values of responding and non-responding patients ($P < 0.05$).

Predicting RT Outcomes with SUV Derived From Bony Aligned Volumes

In contrast to the excellent ability of SUV derived from deformably mapped contours to distinguish responding from non-responding patients, the bony aligned pre-RT volumes overall performed poorly. In the previous ROC analyses, only those metrics that could significantly differentiate responding from non-responding patients were utilized. In the current analysis, no rigidly aligned metrics met this benchmark. In figure 3.10, the ROC curves for post-RT SUV from each technique are shown. For reference, the curve representing performance of pure guessing is plotted as a dashed line. Additionally, the sensitivity and (1 – Specificity) for the post-RT contrast enhanced CT is shown as a single data point for comparison. The post-RT contrast enhanced CT was found to have a sensitivity of 80% and a specificity of 92%.

Decision thresholds that maximize sensitivity and minimize (1 – specificity) for each technique were selected from tabulated ROC curves. For the SUV_{Peak} metric, a post-RT SUV decision threshold of 3.1 was selected from the tabulated ROC analysis. This cutoff point yielded the highest sensitivity of, 100%, and specificity, 94%, for the deformably mapped contours. However, for the rigidly aligned volumes the sensitivity (60%) and specificity (78%) were lower. The decision threshold resulted in 2 false negatives in comparison to zero for the deformed contours. The positive predictive value was 21% and the negative predictive value was 95%. These values were lower than for contrast

enhanced CT, but the positive predictive value was the best for the rigidly aligned PET metrics.

The 50% threshold and GTV contour both performed poorly. For the PET-based threshold contour, a decision threshold of 2.0 was selected. Utilizing this cutoff, the sensitivity was found to be no better than the rigidly aligned SUV_{Peak}, 60%, and the specificity was found to be worse, 61%. One false negative was identified. While the negative predictive value was high, 94%, the positive predictive value was the lowest found, 14%. Results for SUV_{Mean} derived from the rigidly aligned GTV contour were similar. A decision threshold of 1.9 was selected. This resulted in one false negative. The sensitivity was better than all other techniques and identical to contrast enhanced CT, 80%. The specificity was 63%. Furthermore, the positive predictive value of this technique was the second lowest at 18%. Table 3.4 displays the results of all ROC analyses.

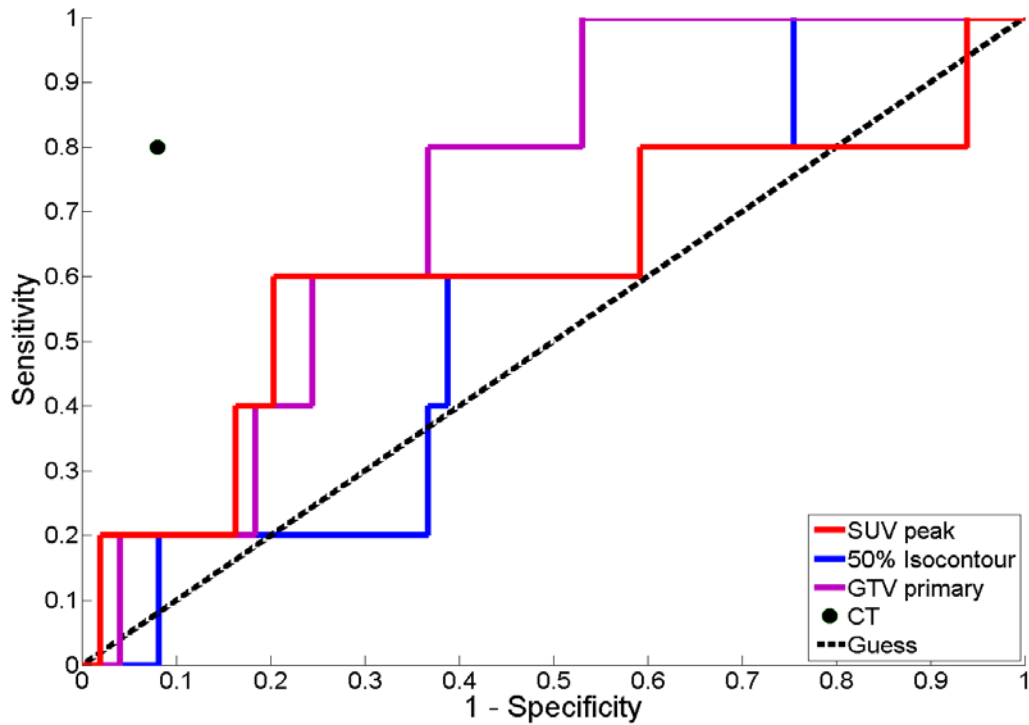


Figure 3.10 ROC curves for rigidly aligned volumes

ROC curves for SUVMean derived from the rigidly aligned GTV, PET-based 50% threshold contour, and SUVPeak are shown. The single data point represents the sensitivity and specificity of contrast-enhanced CT for the study.

	CT	SUV_{Peak}	50% Isocontour	GTV Primary
Cutoff value		3.1	2.0	1.97
True positives	4	3	3	4
False negatives	1	2	2	1
True negatives	45	38	30	31
False positives	4	11	19	18
Sensitivity	0.80	0.60	0.60	0.80
Specificity	0.92	0.78	0.61	0.63
Positive predictive value	0.50	0.21	0.14	0.18
Negative predictive value	0.98	0.95	0.94	0.97

Table 3.4 Accuracy of response prediction with rigidly aligned volumes

Cutoff values were selected from ROC curves to classify patients as normal or abnormal. The raw numbers for each error type are tabulated along with calculated sensitivity, specificity, positive predictive values, and negative predictive values.

Comparing SUV Derived from Bony and Deformable Alignment

To compare the performance of each technique, the area under each curve of the ROC plot was compared pair-wise. SUV_{Mean} derived from the GTV contour resulted in the largest area under the ROC curve, 0.73 (0.55 – 0.90, 95% C.I.). This is in contrast to what was found with the deformably mapped contours, where the GTV contour had the smallest area under the ROC curve. SUV_{Mean} derived from the 50% threshold contour had an area of 0.56 (0.34 – 0.79 95% C.I.). Finally, SUV_{Peak} was found to have an area under the ROC curve equal to 0.62 (0.31 – 0.92 95% C.I.). Comparison of the area under the ROC curves for each rigidly aligned contour found no statistically significant differences ($P>0.05$).

However, when comparing the rigidly aligned contour to the deformably mapped contours the results are markedly different. First, consider the SUV_{Peak} metric. The difference between the areas under each ROC curve was found to be 0.37 (0.04 – 0.71 95% C.I.). Deformably mapping the SUV_{Peak} volume defined on pre-RT PET/CT images to post-RT images was found to significantly increase the area under the ROC curve ($P<0.05$) when compared to bony alignment of this volume. Next, consider the 50% threshold contour. The difference in AUC between the two alignment techniques was found to be 0.37 (0.14 – 0.59 95% C.I.). This increase in AUC is significant at the 95% confidence level. Finally, consider the only contour created on CT images, the GTV. The difference in AUC between the two alignment methods was 0.07 (-

0.07 – 0.22 95% C.I.). This increase in AUC was not significant ($P>0.05$). On the whole, deformable mapping of pre-RT volumes to determine SUV increased the ability of PET/CT to predict response to RT for this data set.

Discussion:

Many studies assessing and predicting response to RT with ^{18}F -FDG PET/CT have reported positive outcomes for patients that showed reduced FDG uptake [89-93]. Quantification of residual disease and subclinical response offers the possibility to obtain a detailed picture of tumor viability and has been sought for some time [94]. Recent studies have confirmed the ability of quantitative PET/CT to distinguish residual disease, but with little benefit over conventional imaging [34]. Limited usefulness of the modality may be related to tumor sampling and recent position papers have suggested that response analyses should move away from the traditionally used single pixel values to region based assessment [95]. However, the presence of gross anatomical changes and multiple imaging studies will require accurate localization of FDG-avid volumes.

The appropriate volume to derive SUV from is a complex question. Total tumor volume was investigated and defined on CT and PET images. One potential benefit of using a larger volume is that they tend to have less variability in SUV than smaller ones [96]. This concept was reinforced by the bony aligned GTV ROC curve. The GTV contour had the largest AUC most likely due to the large volume of tissue contained within, making it less susceptible to incomplete

resolution of the most FDG-avid region. The GTV contour created on the planning CT images was the only metric not able to differentiate the SUV of responding and non-responding patients when deformably mapped to pre- and post-RT images. This may have resulted from the larger volume used to determine SUV (33 cm^3 vs. 1 cm^3) that included necrotic as well as residual cancer cells. Additionally, because the GTV was defined on CT images, misalignment between the inherently aligned PET and CT images could result in calculating SUV for undesired tissue. In one study, misalignment in head and neck PET/CT was found to be as much as 1° or 7 mm [97].

The most widely used, and smallest relative volume, has been SUV_{Max} . The maximum pixel value has been used to assess potential residual disease because it was thought to represent the maximum tumor burden, although it may just represent the most convenient metric. This metric is also known to be greatly affected by image noise and may not be the ideal metric [96]. Alternatively, SUV_{Peak} has begun to gain favor [95]. One potential drawback of the SUV_{Peak} method would be a non-centrally located maximum pixel. In this case, the 1 cm^3 could potentially include adjacent normal tissue. This scenario could be quite troubling for tumors with maximal dimensions less than 1.2 cm. The average lesion size for the cohort as measured on pre-RT CT was 2.7 cm. Lesions less than 1.2 cm did not have SUV_{Peak} determined in this study. The average largest diameter measureable on post-RT CT was 2.9 cm.

For the current cohort, SUV_{Peak} of non-responding patients was significantly higher than patients who responded to treatment (4.6 ± 1.3 vs. 2.4 ± 0.7 , $P < 0.01$) when derived from the deformably mapped volume. After a decision threshold was selected, SUV_{Peak} was the only method to have zero false negatives. Additionally, the PPV for residual disease was 63%. For post-RT contrast enhanced CT, PPV was only 50%. Because SUV_{Peak} is centered on the most metabolically active region of a tumor, viable residual cancer cells are likely to be circumscribed by this volume [98-100]. Assessing response to treatment based upon FDG-uptake in the total tumor volume may reduce the post-RT SUV and mask the metabolic activity in this more critically important tumor region.

Although significant differences were found between absolute post-RT SUV values of responding and non-responding patients this separation may be artificial. The SUV value can be affected by many factors including patient preparation, incorrect cross-calibration of scanner and dose calibrator, and variable injection to imaging times [101]. Assessing the SUV of background tissue can be a powerful discriminator in determining if differences between two patient groups are real. The liver has been suggested as a stable background measurement point, but all patients in this cohort were head and neck cancer patients and PET/CT images did not include the liver [95]. Alternatively, non-visual cortex brain tissue was shown in one study to function well as a background measurement point [102]. The brainstem was selected as a background measurement point because it is often contoured in head and neck

treatment planning. Post-RT SUV_{Mean} for the brainstem in the responding group (4.7 ± 1.1) and non-responding group (3.9 ± 1.0) were not significantly different ($P>0.05$). This strengthens the case for SUV differences between the two patient groups representing real effects.

When comparing the deformably aligned volumes to the alternate alignment to a bony structure the advantage was clear. The alternative method of alignment was based upon registration to the C2 vertebra. For all volumes considered, SUV derived from bony aligned volumes was not significantly different between responding and non-responding patients ($P>0.05$). Additionally, the use of deformably mapped volumes was found to significantly increase the AUC for the SUV_{Peak} volume and the threshold contour ROC curves, but not for the CT-based GTV.

Conclusions:

In conclusion, deformable image registration was found to improve region based response analysis. Specifically, for the most metabolically active tumor regions, mapping this volume across longitudinal studies significantly improved the predictive power of PET/CT over bony alignment ($P<0.05$). A post-RT SUV decision threshold of 3.5 was selected for SUV_{Peak} . Utilizing this cutoff, the PPV for PET/CT was 63%, markedly higher than post-RT contrast enhanced CT (50%). However, because the number of non-responding patients was small ($n = 6$), the results of the study need validation. The results of this aim support the hypothesis that voxel-by-voxel mapping of primary tumor volumes to serial F-

FDG PET/CT images would improve region based response analysis. Incorporating deformable image registration into quantitative treatment response monitoring studies to more accurately localize residual disease may greatly impact patient management and allow for more timely intervention.

Chapter 4

Specific Aim 2a

Evaluate the potential to determine normal organ-at-risk (OAR) toxicity in patients treated with intensity modulated radiotherapy (IMRT) for head and neck cancers with ^{18}F -FDG PET/CT.

. *Working Hypothesis: Parotid gland SUV will function as an objective imaging biomarker of salivary toxicity correlating with stimulated salivary flow ($P<0.05$).*

Purpose

As described in Chapter 2, current methods used to assess salivary toxicity in patients treated for head and neck cancers with RT are limited and perform poorly. For instance, for the cohort employed in specific aim 1 only 30/107 screened patients had any measure of salivary toxicity collected as part of the routine standard of care. When salivary toxicity measures were collected, they were limited to xerostomia questionnaires only collected after RT. Without baseline measures, assessing the toxicity of a treatment is particularly challenging.

The potential of imaging biomarkers to measure the effects of cancer treatments have been well documented [92, 103, 104]. Arming clinicians with an

in vivo continuous variable that functions as a surrogate for clinical endpoints has powerful implications for patient management. Yet, relatively little attention has been paid to investigating the ability of imaging biomarkers to quantify normal organ-at-risk (OAR) toxicity [105-107]. The near drought of work in this area has lead to the establishment of groups like the Image Response Assessment Team (IRAT) whose sole purpose is to further the use of quantitative imaging for treatment response to therapy in the clinic. Additionally, recent changes in reimbursement guidelines for PET from the Centers of Medicare and Medicaid will only increase the number of number of initial staging and follow-up PET scans being performed [108].

Because of the expanding role of PET/CT in patient management the opportunity exists for uses tangential to initial staging or follow-up. Specifically, for patients with head and neck cancer, incidental collection of parotid ^{18}F -FDG uptake was investigated as a surrogate for salivary function. Sialometric data collected on the same day as PET/CT imaging studies were correlated with SUV determined for the parotid glands in this study. Utilization of imaging biomarkers in this manner may allow for patient specific refinement of radiation treatment plans to reduce normal OAR toxicity.

Methods and Materials

Patient Cohort

To test the working hypothesis for specific aim 2, a separate cohort was employed in which patients had pre- and post-RT PET/CT imaging and measures of salivary function collected at the same time point. Patients were selected from an ongoing IRB approved protocol (LAB07-0050) designed to establish a database of the effects of RT on salivary flow. Inclusion criteria included patients with histological confirmation of head and neck cancer set to receive definitive RT, concurrent chemotherapy, or RT after surgery who were at least 18 years of age. Patients who had previous RT of the head and neck were excluded. Sialometric as well as subjective measures of salivary toxicity were collected at five time points. The first collection was prior to the start of treatment. While only a snapshot of salivary gland function, this time point will serve as the baseline measurement for saliva production. The second time point is following RT, approximately 6 weeks. The third time point is at four to six months after the completion of RT. The fourth collection date is at one year post-RT (± 2 months). Finally, the last collection time point is two years after the completion of treatment (± 3 months).

Sialometric Evaluation

Whole mouth stimulated and unstimulated saliva was collected at each time point. For the unstimulated collection, patients were instructed to refrain

from eating, drinking, or dental hygiene for at least 30 minutes. The goal was to minimize the impact of all oral stimuli. The subject was instructed not to induce salivation. Each collection period lasted five minutes; this was shown on a digital timer in view of the patient. During collection patients were instructed not to swallow and let the saliva accumulate in the floor of the mouth. Every sixty seconds the patient was instructed to expectorate the accumulated saliva into 100 mL vials. This procedure was repeated every minute for the five minute collection period. Each empty vial was massed prior to saliva collection. After the patient's saliva was collected the vial and saliva were massed. The difference between the two measurements is the mass of the saliva.

In order to measure the stimulated whole mouth saliva, the patient was instructed to rest for five minutes after the collection of unstimulated whole mouth saliva. To induce salivation, 20 mL of a citric acid solution was held in the patient's mouth for one minute. The same methodology described for the unstimulated salivary collection was employed to collect stimulated salivary mass.

For each patient stimulated and unstimulated baseline salivary flow measurements were assessed for abnormality. Unstimulated whole salivary flow rates less than 0.1 mL/min and stimulated whole salivary flow rates less than 0.5 mL/min have historically been labeled as abnormal [109]. The salivary flow thresholds for abnormal flow were converted to mass cutoffs for the collection period of five minutes utilized in this protocol. For the unstimulated

sialometric data, whole mouth salivary mass less than 0.5 g was classified as abnormal. For the stimulated salivary data, whole mouth salivary mass less than 2.5 g was defined as abnormal baseline salivary function.

Subjective Evaluation

Patient self-reported xerostomia questionnaires were utilized for subjective evaluation of salivary toxicity [75]. The patient completed a set of eight questions on the degree of difficulty talking, chewing, and swallowing due to perceived dryness. Each question received a score of 1 – 10, with greater numerical value corresponding to greater perceived complication. Finally, a summary score was created by summing the results of each question to produce a value between 0 and 80.

Xerostomia Grade

Salivary toxicity was assessed using the EORTC/RTOG late effects toxicity scoring subjective, objective, management, analytic (SOMA) scale. The analytic scale was employed to classify patients into grade 0 – 4 xerostomia at the time of first follow-up (approximately 50 days post RT). The analytic scale separates patients into grades by assessing the saliva produced as compared to baseline. The percentages are presented in table 4.1.

	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Salivary flow/quantity	> 95% of pre-treatment	76% - 95% of pre-treatment	51% - 75% of pre-treatment	26% - 50% of pre-treatment	0% - 25% of pre-treatment

Table 4.1 RTOG/EORTC late effects analytic scale

Salivary toxicity grades are defined by the relative change in saliva production.

Volume Delineation

In order to evaluate the potential of using SUV as an objective imaging biomarker of salivary function, physician created parotid contours created on planning CT images were utilized. The general delineation guidelines for the human parotid glands follow specific anatomical landmarks [110]. The superior boundary extends to the external auditory canal and mastoid process, figure 4.1a. Inferiorly, the parotids were contoured to the submandibular space, figure 4.1b. In the anterior direction, masseter muscle forms the boundary. This is shown in figure 4.1c. The sternocleidomastoid muscle forms the posterior border. Laterally, the parotids are bounded by subcutaneous fat. Finally, the medial border is formed by the posterior belly of the digastric muscle.

However, these are only guidelines. For instance, in 20% of the cases, the parotid will extend anteriorly past the masseter muscle [111]. Other, non-patient, variations will also exist in the parotid contours. For example, the adaptive protocol cohort had contour created by multiple physicians ($n = 7$). In a 2005 study of the inter-observer contouring variations for head and neck anatomy, it was found that when compared to a mixed volunteer cohort, experienced head and neck radiation oncology specialist were able to reduce volume coefficient of variation (defined as the quotient of the standard deviation and the mean volume) by 16% for the parotid glands. The presence of dental artifacts only impacted the contour volume coefficient of variation by 6%. The center-of-volume standard deviation was reduced to 1 mm [112]. These

variations were found to be acceptable. It is understood that while both anatomical and observer variations exist in the parotid contours, this is not expected to fundamentally alter the conclusions of the technique.

The parotid glands were originally contoured on planning CT images collected for patient simulation. These physician created volumes were then deformably mapped to the pre- and post-RT PET/CT images using the enhanced demons algorithm and the same methodology described for specific aim 1. In figure 4.2, an example patient's pre-RT PET/CT with deformably aligned parotid contours is shown. In figure 4.2b and 4.2d, FDG-avid parotid glands are shown before and after RT, respectively. The PET image has the upper window level set to the maximum SUV value within the parotids with the lower window set at 42% of this value [113].

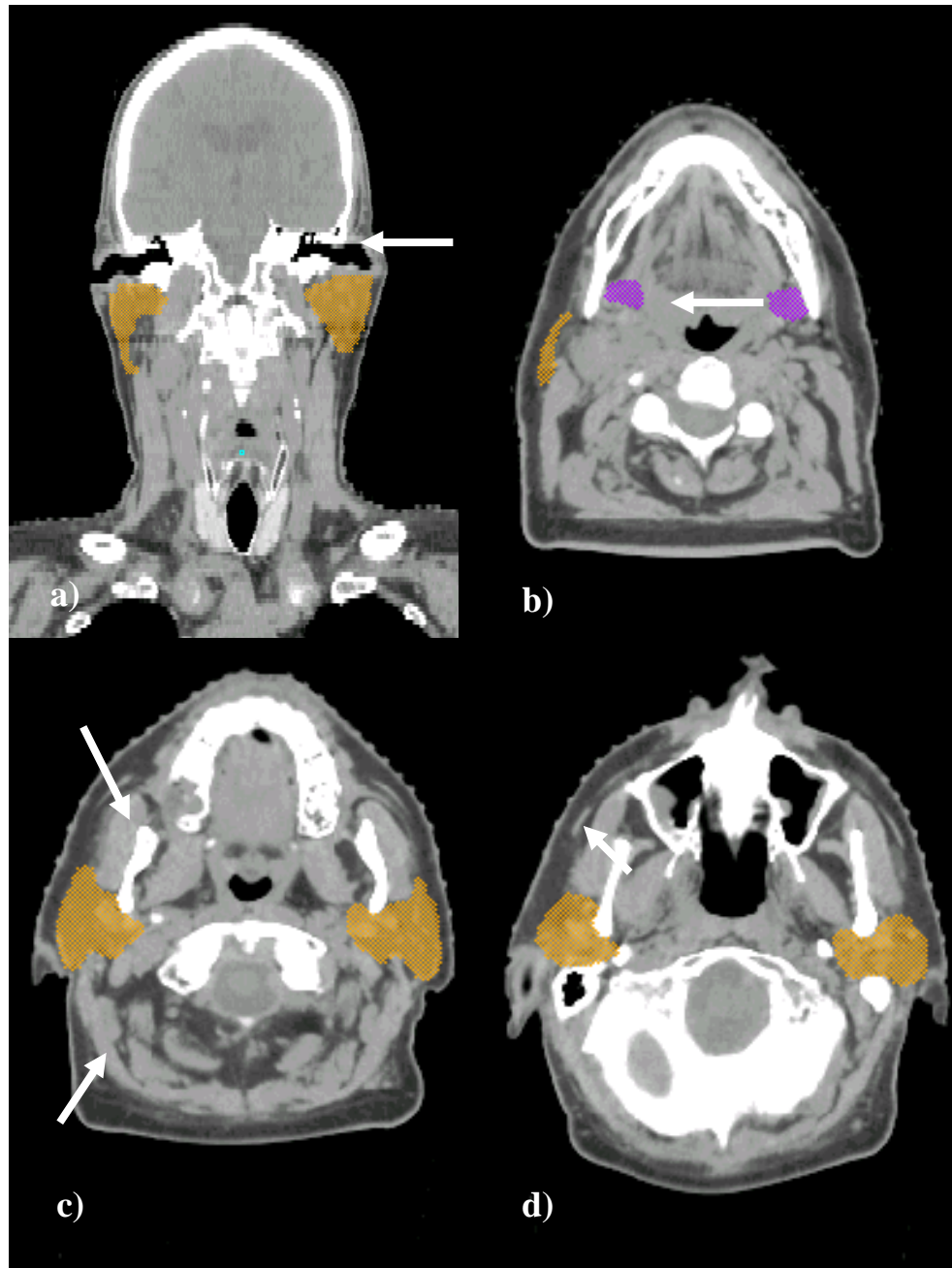


Figure 4.1 Parotid anatomical boundaries

The anatomical structures that form the common boundaries for the parotid glands (orange) are depicted in each frame (white arrow). The submandibular glands are shown in purple. **(a)** The external auditory canal forms the superior boundary of the parotid glands. The inferior boundary is formed by the submandibular space **(b)**. In **(c)**, the masseter muscle forms the anterior boundary with the sternocleidomastoid muscle forming the posterior border. The platysma forms the lateral border **(d)** while the posterior belly of the digastric muscle forms the medial border.

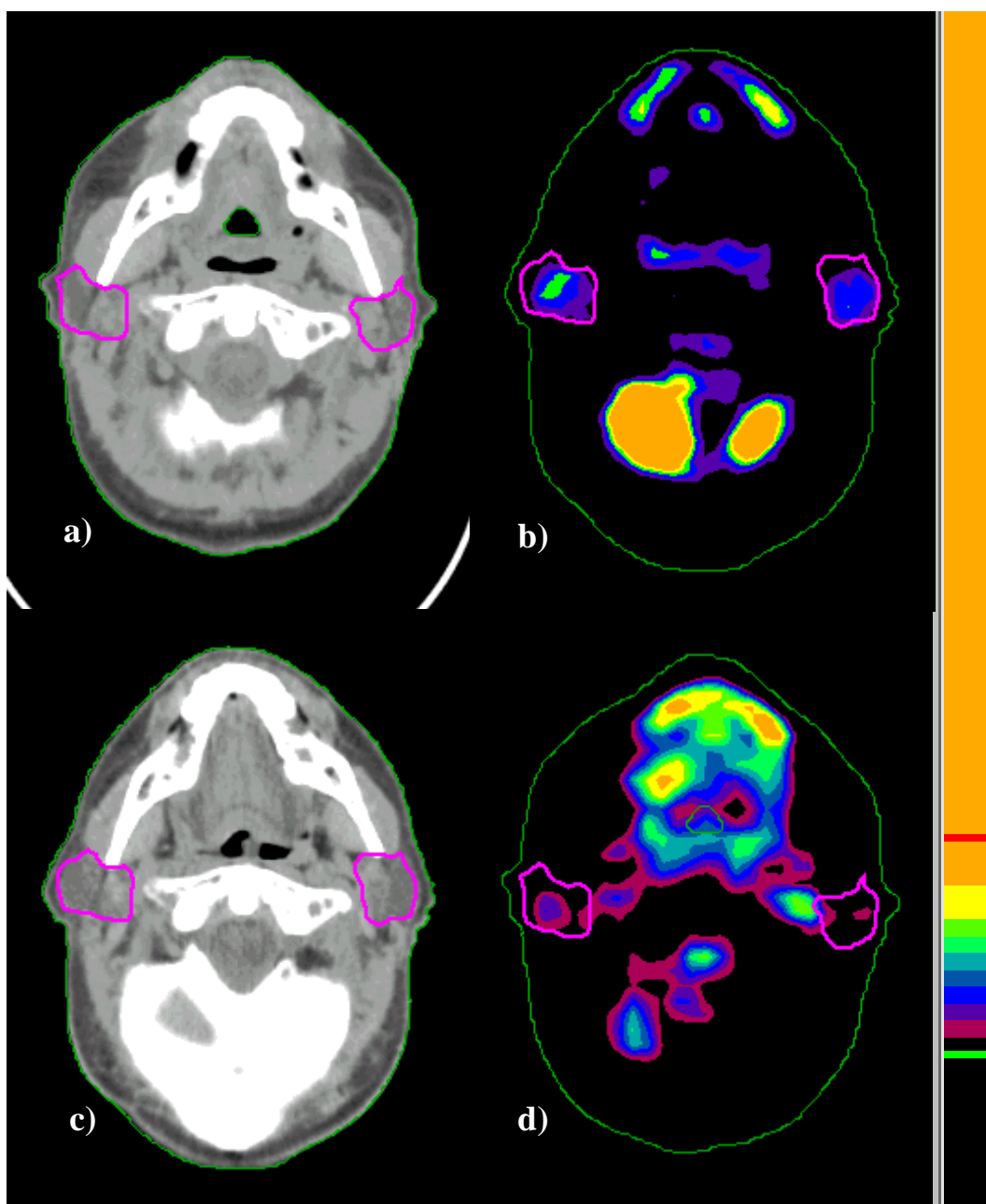


Figure 4.2

Pre-RT CT (a) and PET (b) co-registered images with deformably aligned parotid contours (purple). The green contour denotes the skin surface. The post-RT CT (c) and PET (d) are also shown.

Results:

Patient Characteristics

Fourteen patients were available for the current analysis that had pre- and post-RT PET/CT as well as salivary evaluation performed. The cohort was composed mostly of males, $n = 12$, and few females, $n = 2$. The average age of the group was 57 years with a range of 49 – 75 years. Normal variations in saliva production have been shown to be age and gender independent [114, 115].

Tumor staging was based upon the American Joint Committee on Cancer. Nine patients had stage four, three patients had stage three, and one patient each had stage one or two disease. All primary tumors were located in the oropharynx. The prescription dose was 70 Gy, $n = 12$, and 66 Gy, $n = 2$. Patients were treated once daily with a fraction size of 2.0 Gy, $n = 4$, 2.12 Gy, $n = 8$, or 2.2 Gy, $n = 2$. The mean dose delivered to the parotid glands was 28 ± 6 Gy. The ipsilateral parotid gland, defined as the most proximal of the paired glands to the primary tumor, received mean doses of 36 ± 12 Gy. The contralateral parotid gland received considerably lower mean doses of 20 ± 2 Gy.

Baseline Saliva Output

Measurements of baseline salivary flow were used to establish patients with pre-existing abnormal salivary function. For the fourteen accrued patients,

unstimulated saliva production for a five minute collection period had an average value of 1.92 ± 1.04 g with a range of 0.45 – 4.04 g. The distribution of baseline unstimulated flow is illustrated in figure 4.3. Utilizing the specified cutoff for abnormal salivary function, for unstimulated salivary flow, 0.5 g, one patient was identified as having abnormal baseline function. The vertical reference line in figure 4.3 depicts this cutoff.

For the baseline stimulated saliva production, an average mass of 5.51 ± 2.4 g was found with a range of 1.51 – 9.46). The histogram in figure 4.4 graphically displays the distribution of baseline stimulated saliva production. The cutoff for abnormal baseline stimulated whole mouth salivary flow, 2.5 g, identified one additional patient. Furthermore, because xerostomia is fundamentally a subjective disorder, baseline XQL questionnaire scores were utilized to identify patients who believed their function was abnormal. Two patients were found with XQL summary scores above the 40 threshold. These data are illustrated in figure 4.5. The scale XQL scale (0 – 80) has been truncated in the figure.

In summary, 4/14 patients were identified as having abnormal baseline salivary function through objective and subjective measures. The average unstimulated saliva mass in the abnormal baseline function group was 1.67 ± 0.89 g. The range of salivary mass was 0.67 – 2.84 g. The normal baseline function group had an average unstimulated saliva mass of 2.02 ± 1.12 g with a range of 0.45 – 4.04 g. There was no significant difference found between the

two groups ($P>0.05$). Even when the two patients classified as abnormal based upon baseline XQL summary scores were not considered, the unstimulated saliva mass was not significantly greater in the “normal” group. For the simulated salivary flow, the abnormal baseline function group produced an average of 4.32 ± 2.97 g with a range of 1.51 – 7.80 g. The normal baseline stimulated flow group produced an average saliva mass of 5.98 ± 2.19 g. The range of saliva mass values was 3.12 – 9.46 g. No significant difference was found between the stimulated saliva mass of normal and abnormal baseline salivary function patients ($P>0.05$). However, if the patients classified as abnormal based upon XQL summary scores are not considered the stimulated saliva mass for the normal function group is significantly greater than the abnormal function group, 1.9 ± 0.5 vs. 6.1 ± 1.1 $P<0.05$.

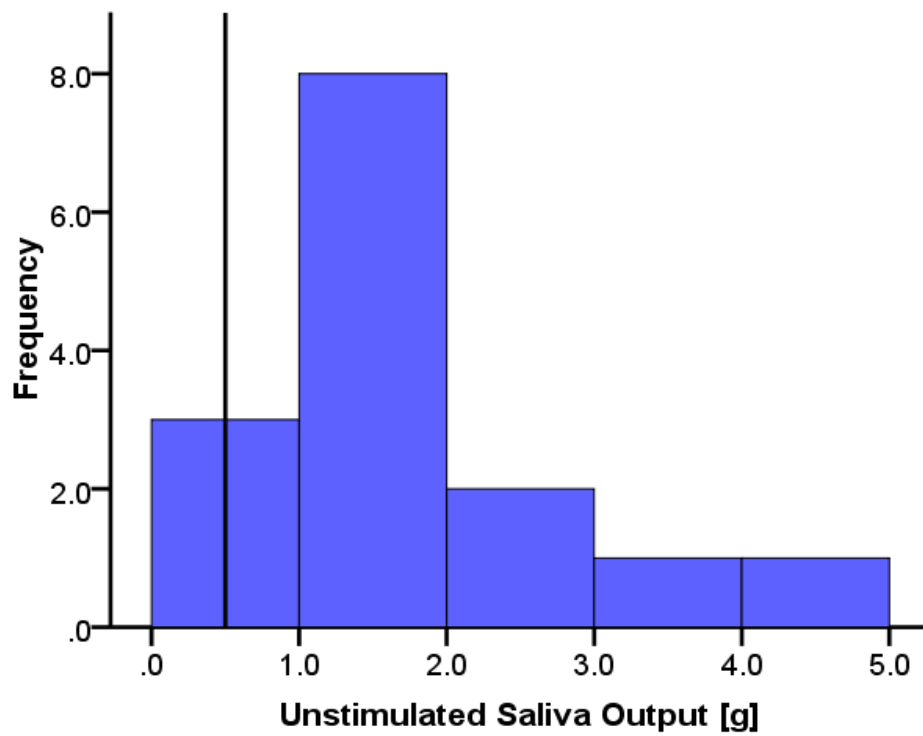


Figure 4.3 Baseline salivary mass

Unstimulated salivary mass collected prior to treatment. The vertical line represents the cutoff for “abnormal” function.

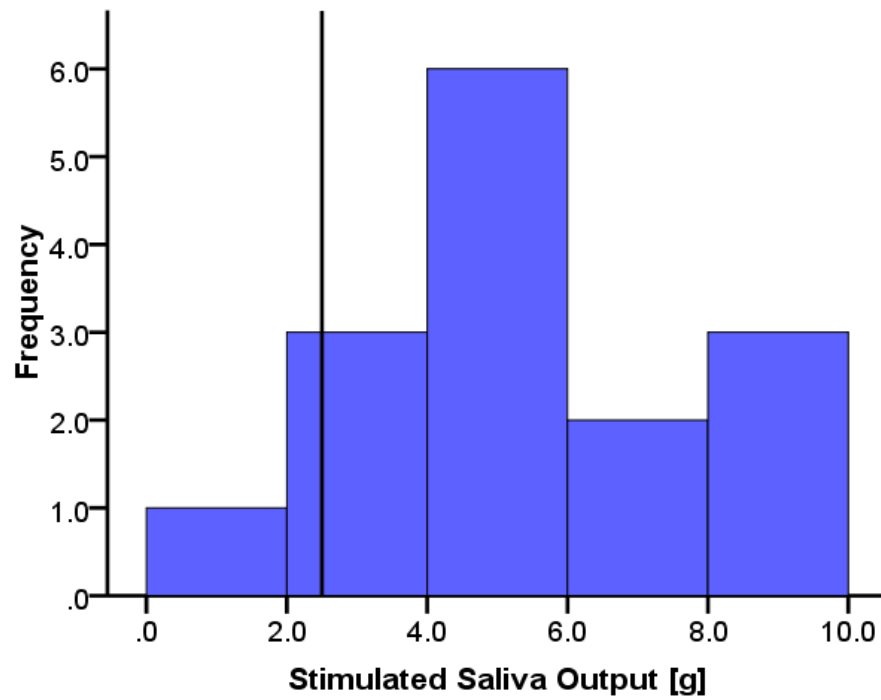


Figure 4.4 Baseline salivary mass

Stimulated salivary mass collected prior to treatment. The vertical line represents the cutoff for “abnormal” function.

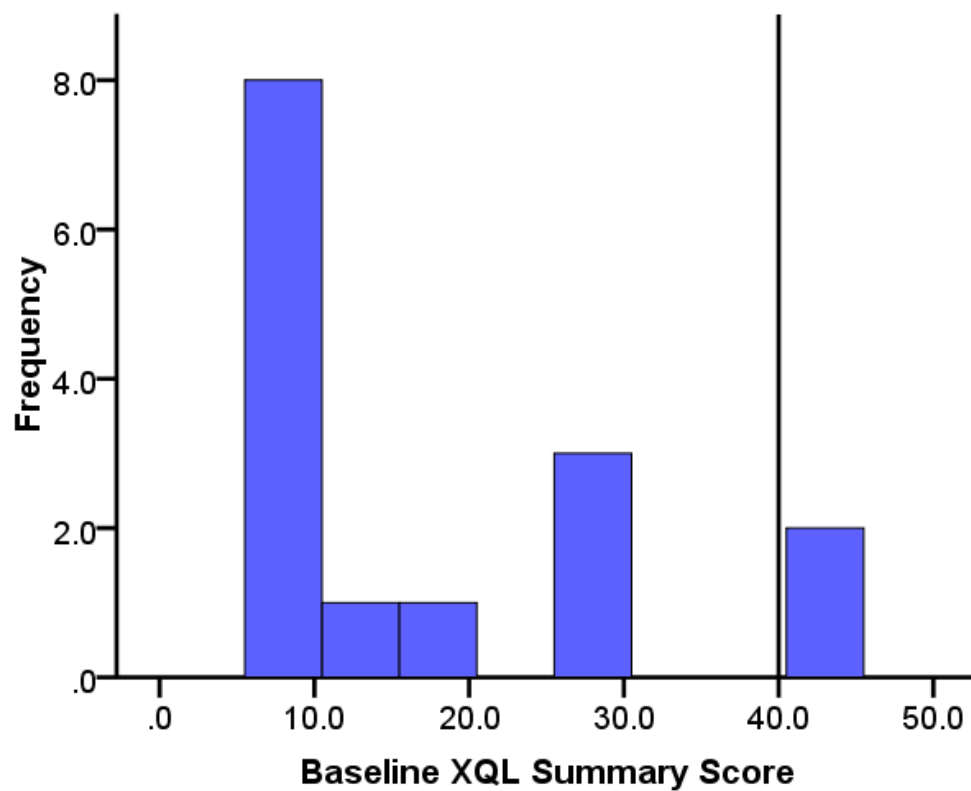


Figure 4.5 Patient perceived baseline salivary function

XQL summary score distribution prior to treatment. The vertical line represents the median of the scale. Note XQL scale (0 – 80) is truncated for display purposes.

Correlating Baseline Salivary Function and SUV

Following the deformation of all physician contoured pre-RT parotid volumes to the pre-RT PET/CT images, SUV was determined for the parotid glands. Parotid gland pre-RT SUV for the baseline abnormal patients was found to be (1.2 ± 0.2) , $n = 4$. For the normal baseline salivary function patients, parotid gland pre-RT SUV was found to be (1.6 ± 0.3) , $n = 10$. This difference in SUV between the normal and abnormal baseline salivary function patients was significant ($P < 0.05$). This was in contrast to the insignificant differences of the stimulated and unstimulated saliva mass between the two patient groups. The distribution of pre-RT SUV is shown in figure 4.6.

Next, the correlation between produced saliva and pre-RT SUV was investigated. In figure 4.7, the relationship between stimulated saliva mass and pre-RT SUV is shown. Moderate positive correlation between the two metrics, Pearson correlation = 0.41, was not significant ($P > 0.05$). For the unstimulated saliva mass produced, stronger negative correlations were observed, Pearson correlation = -0.57, that were insignificant ($P > 0.05$). Finally, the relationship between the subjective XQL summary score and SUV were assessed. This is shown in figure 4.8. Spearman's rho showed moderate negative correlation between the subjective metric and pre-RT SUV, $\rho = -0.46$, $P > 0.05$.

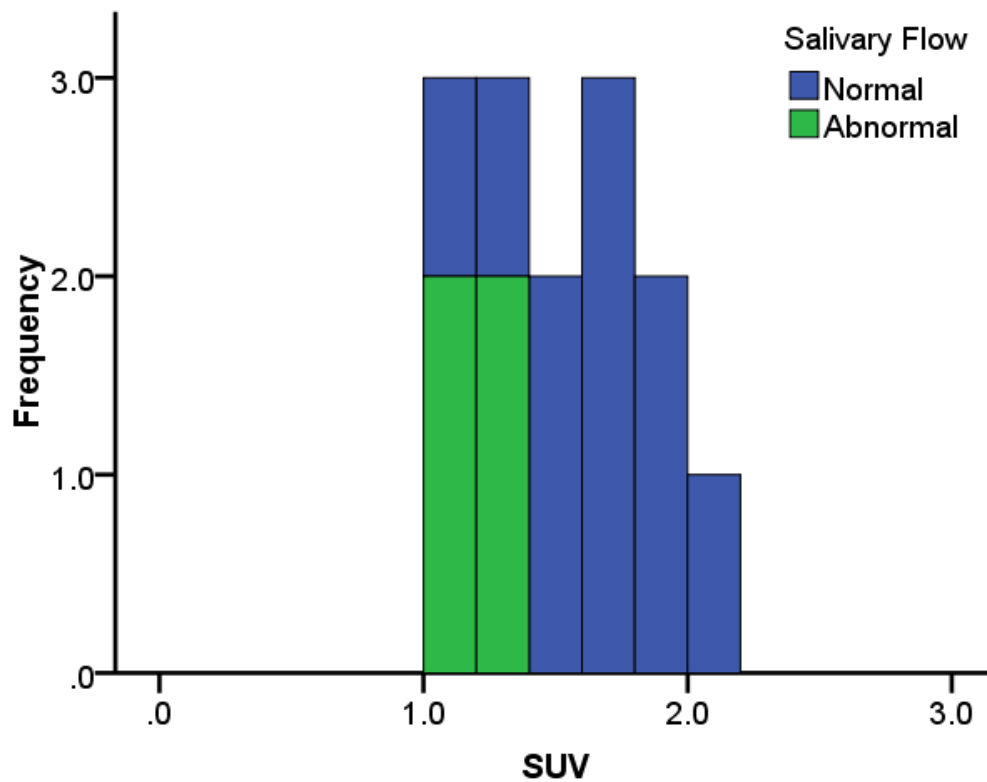


Figure 4.6 Baseline parotid SUV distributions

The histogram distribution of parotid SUV values for the “normal” (blue) and “abnormal” (green) baseline salivary function patients.

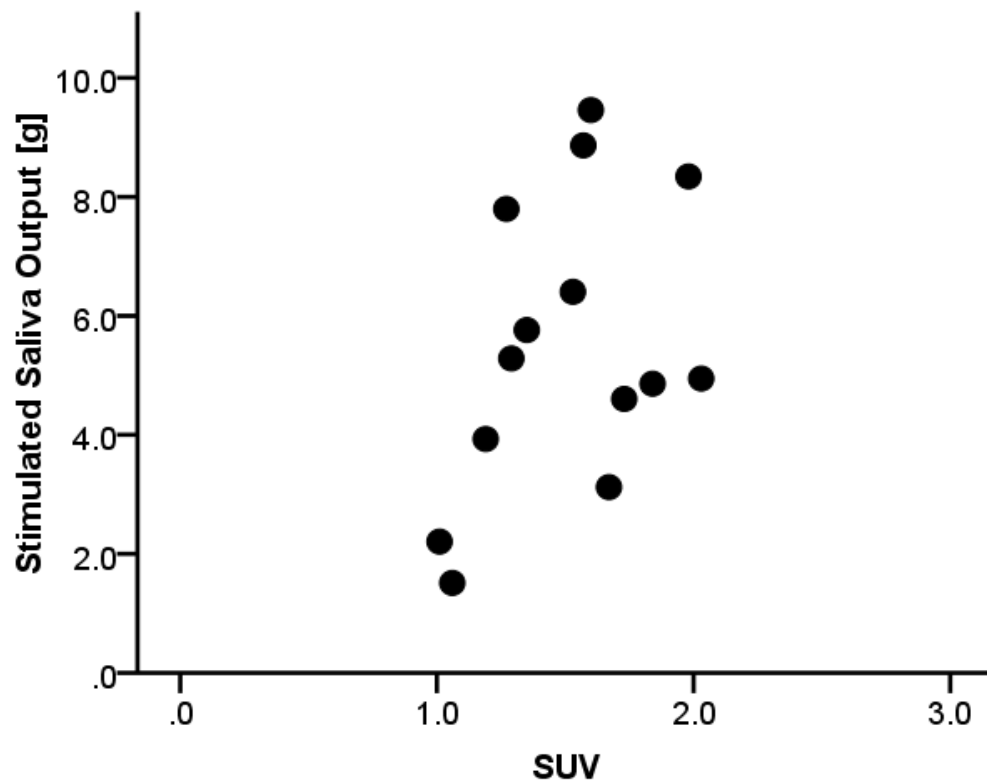


Figure 4.7 Stimulated saliva mass versus SUV

The baseline stimulated saliva mass [g] is plotted as a function of baseline parotid SUV. Pearson correlation = 0.41, $P > 0.05$.

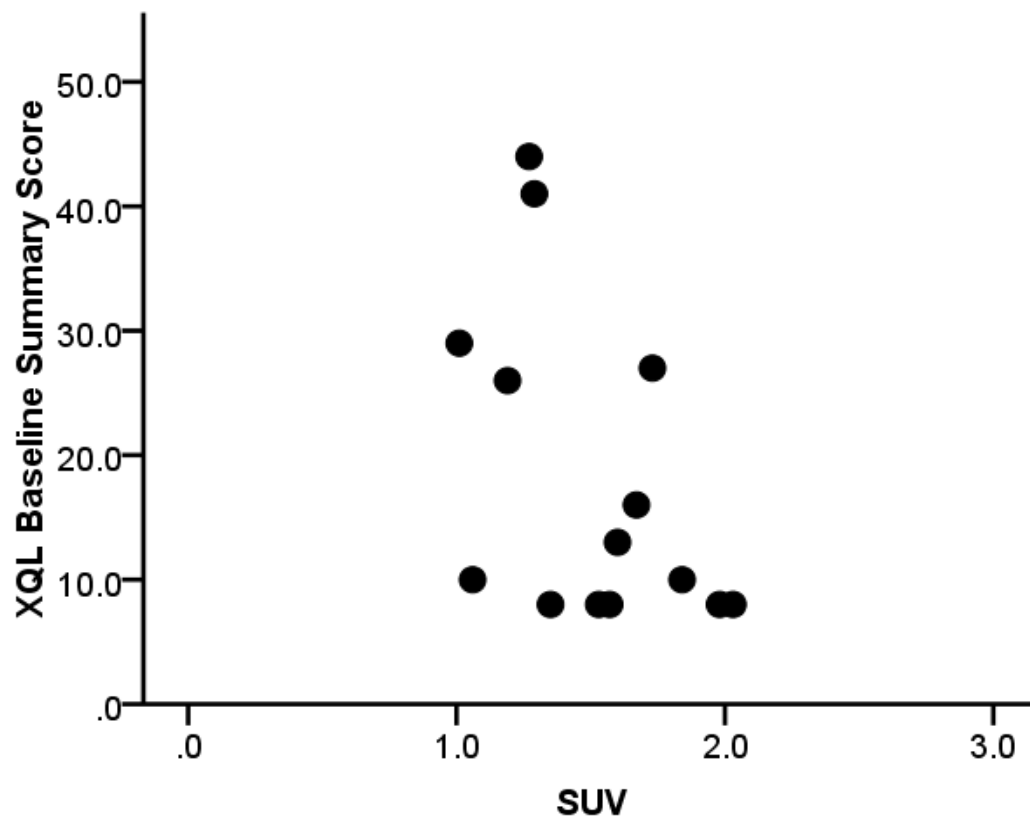


Figure 4.8 Baseline XQL score versus parotid SUV

The pre-treatment XQL summary scores did not show a strong correlation to baseline parotid SUV.

Post-RT Salivary Output

Measurements of post-RT saliva production were utilized to establish salivary toxicity. The average time to first follow-up saliva measurement was 52 days after the completion of RT (range 42 – 62 days). Although fourteen patients were analyzed for baseline correlations of saliva mass and SUV, only eight were employed for the post-RT analysis. The data set was restricted for several reasons. First, two patients failed to have first follow-up saliva collected. Second, patients previously identified as having abnormal baseline salivary function were not included in the analysis, $n = 4$. In this manner, potential confounding factors associated with the deviant salivary function were eliminated.

The average unstimulated saliva mass produced after RT was 1.16 ± 1.52 g with a range of 0.47 – 4.90 g. The average unstimulated saliva mass produced was reduced by 49% compared to baseline masses. The distribution of post-RT unstimulated saliva collected at the first follow-up is shown in figure 4.9. For the stimulated condition, produced saliva had an average mass of 2.26 ± 2.08 g. The range of simulated mass was 1.00 – 7.20 g. The stimulated saliva distribution is shown in figure 4.10. The average stimulated saliva mass produced was reduced by 50% when compared to baseline. The distribution of fractional stimulated saliva output is shown in figure 4.11. Based upon the EORTC/RTOG Late Effects Analytic Scale, patients with grade 4 ($n = 2$), grade 3

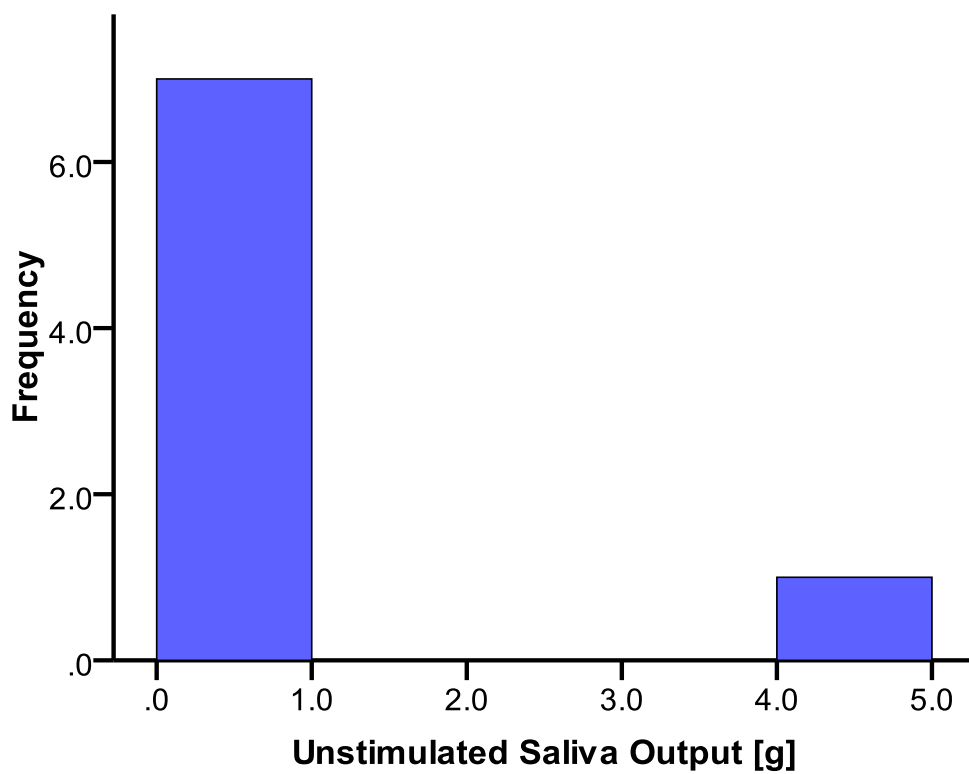


Figure 4.9 Histogram distribution of post-RT unstimulated salivary mass

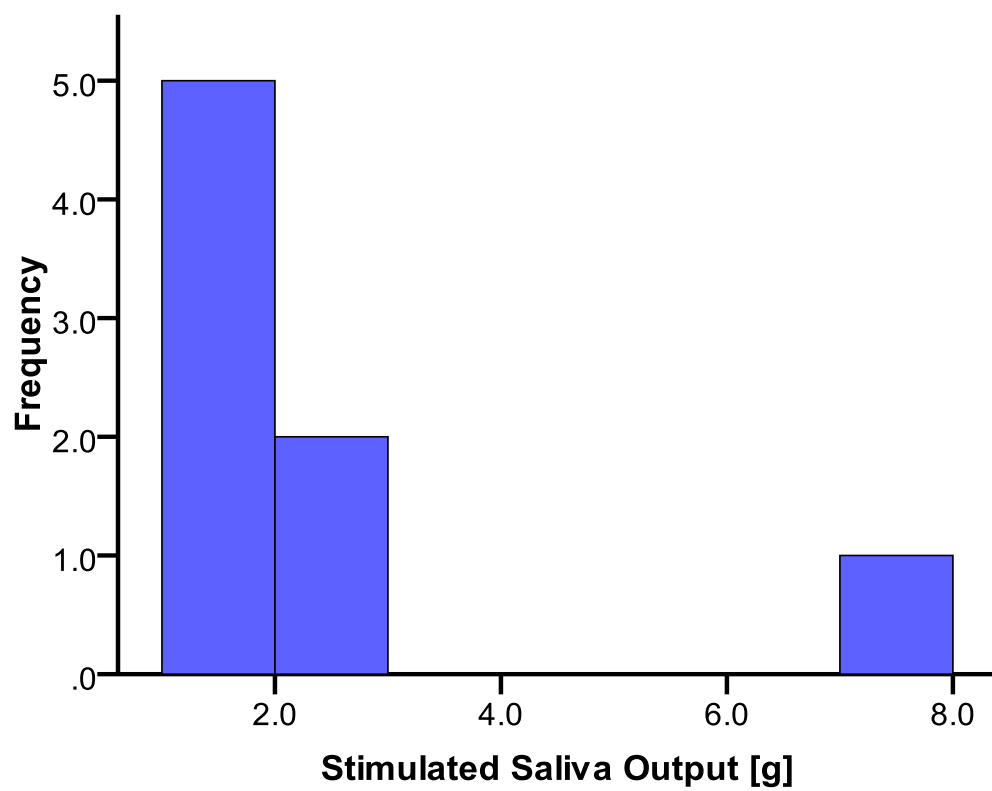


Figure 4.10 Histogram distribution of post-RT stimulated salivary mass

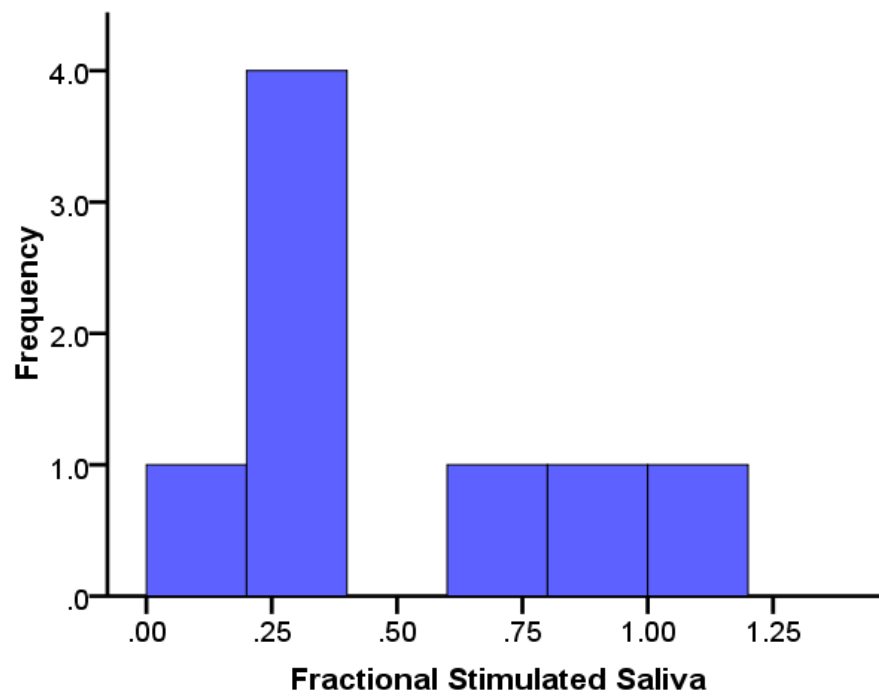


Figure 4.11 Histogram distribution of fractional stimulated saliva mass

(n = 3), grade 2 (n = 0), grade 1 (n = 2) and grade 0 (n = 1) xerostomia were identified.

Correlating Post-RT Salivary Output and SUV

After the pre-RT parotid volumes were deformably mapped to the post-RT PET/CT images, post-RT SUV was collected for the parotid glands and correlated with post-RT salivary output measures. In figure 4.12, unstimulated salivary flow versus absolute post-RT SUV of the parotid glands is plotted. Apart from one easily identifiable patient, all unstimulated salivary output was less than 1 g for the five minute collection period. Post-RT SUV for the cohort ranged from 1.0 – 1.5. Contrast this with the pre-RT unstimulated salivary output where 2/3 of the patients with unstimulated flow less than 1 g were classified as having abnormal baseline flow. Interestingly, the patient with the greatest post-RT unstimulated salivary mass (4.9 g) also had the greatest pre-RT unstimulated mass (4.0 g). Because the parotid glands do not produce the majority of the unstimulated salivary output, correlations between SUV and unstimulated saliva mass were not expected to be strong nor significant. This expectation was confirmed (Spearman's $\rho = 0.45$, $P > 0.05$).

Next, the stimulated salivary output was investigated. In figure 4.13, the stimulated saliva mass versus absolute post-RT parotid SUV is plotted. As was observed for the unstimulated saliva, the stimulated saliva mass range (1.0 – 7.2 g) occupied a small range of SUV values (1.0 – 1.5). The stimulated salivary mass was expected to correlate much closer with parotid SUV. However, as

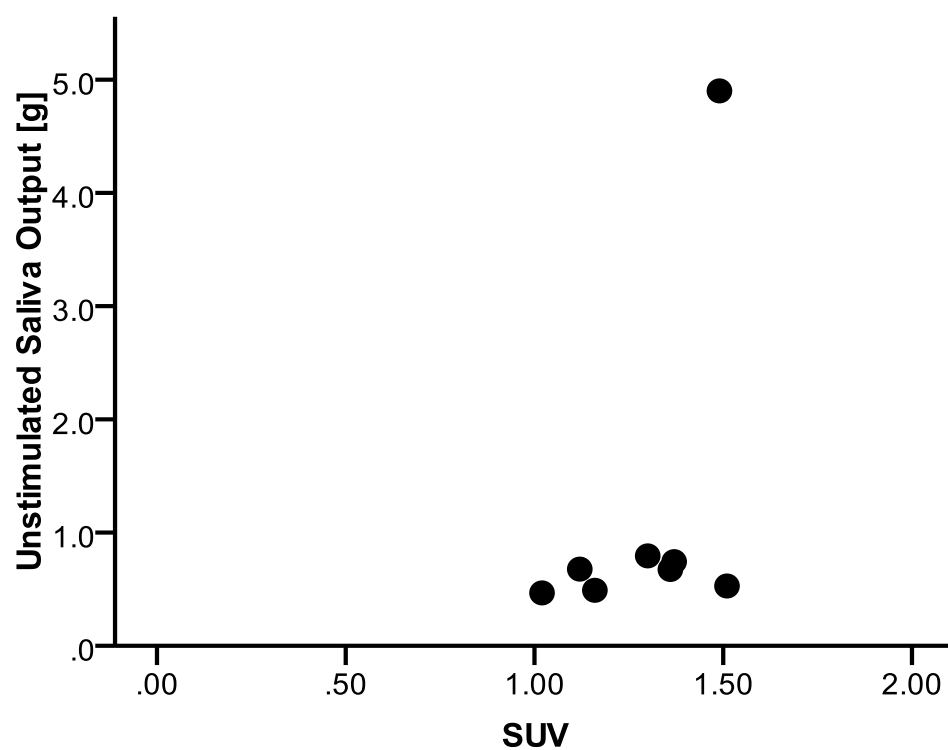


Figure 4.12 Unstimulated saliva mass and SUV

The absolute post-RT unstimulated saliva mass did not correlate strongly with post-RT parotid SUV. Spearman's $\rho = 0.45$, $P > 0.05$

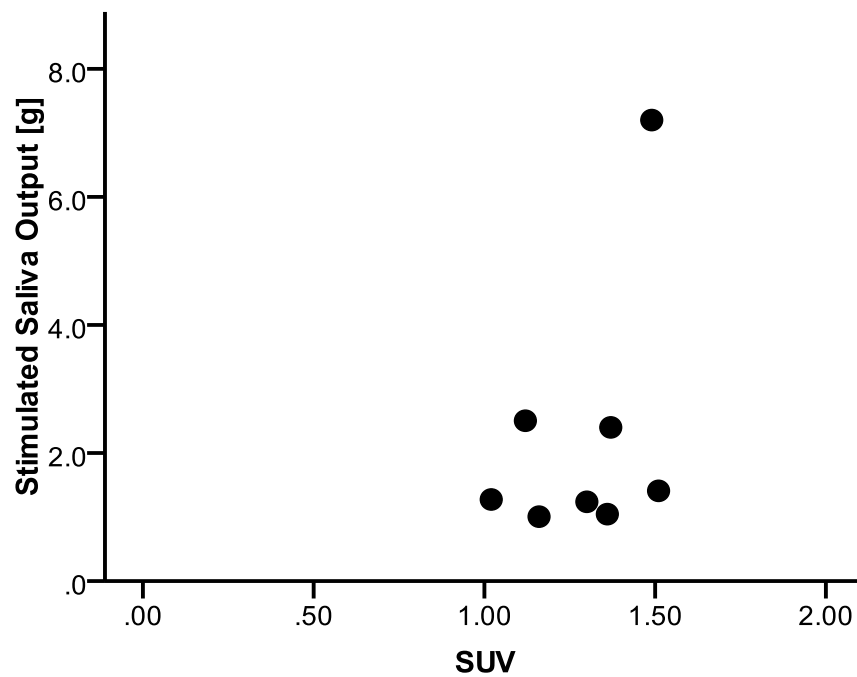


Figure 4.13 Stimulated saliva mass and SUV

The absolute post-RT stimulated saliva mass did not correlate strongly with post-RT parotid SUV. Spearman's $\rho = 0.31$, $P > 0.05$

was observed for baseline stimulated salivary mass and absolute pre-RT SUV, minimal correlation was observed (Spearman's $\rho = 0.31$, $P > 0.05$).

Additionally, the correlation between subjective measures of post-RT salivary toxicity and SUV were assessed. The relationship between the subjective metric and SUV is illustrated in figure 4.14. As expected XQL summary score was negatively correlated with SUV, although the results were not significant ($P > 0.05$).

Finally, relative changes in stimulated salivary output and SUV were investigated. For 7/8 patients in the cohort, stimulated saliva output at the time of follow-up imaging was reduced. Fractional stimulated saliva output for this group was 0.41 ± 0.28 . Reduction of saliva output paralleled SUV reduction; the imaging biomarker value was 0.80 ± 0.14 . One patient's biomarker value indicated increased SUV from baseline (1.11). This patient had a corresponding increase in stimulated saliva output with a fractional stimulated saliva value of 1.14. These metrics demonstrated strong and significant positive correlation (Pearson correlation = 0.93, $P < 0.01$). The relation is illustrated in figure 4.15. Furthermore, fractional SUV was correlated with xerostomia grade based on the RTOG/EORTC late effects analytic scale. Xerostomia grade was negatively correlated with fractional SUV (Spearman's $\rho = -0.964$, $P < 0.01$). This relation is shown in figure 4.16.

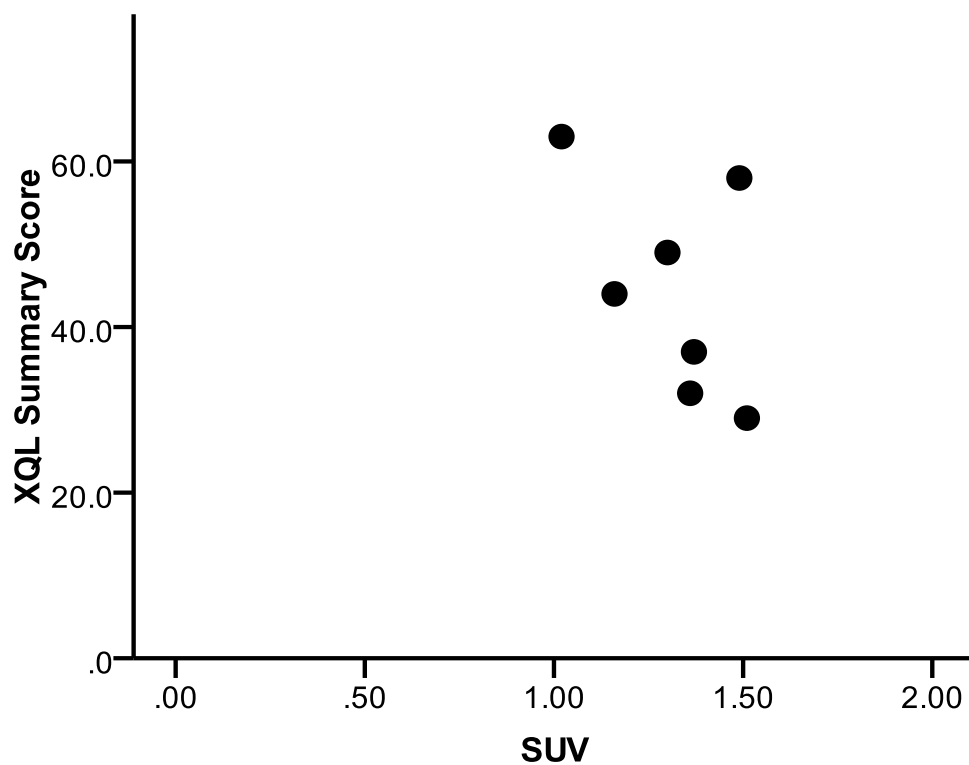


Figure 4.14 Post-RT XQL score and SUV

The post-RT XQL summary score did not correlate strongly with post-RT parotid SUV. Spearman's $\rho = -0.56$, $P > 0.05$

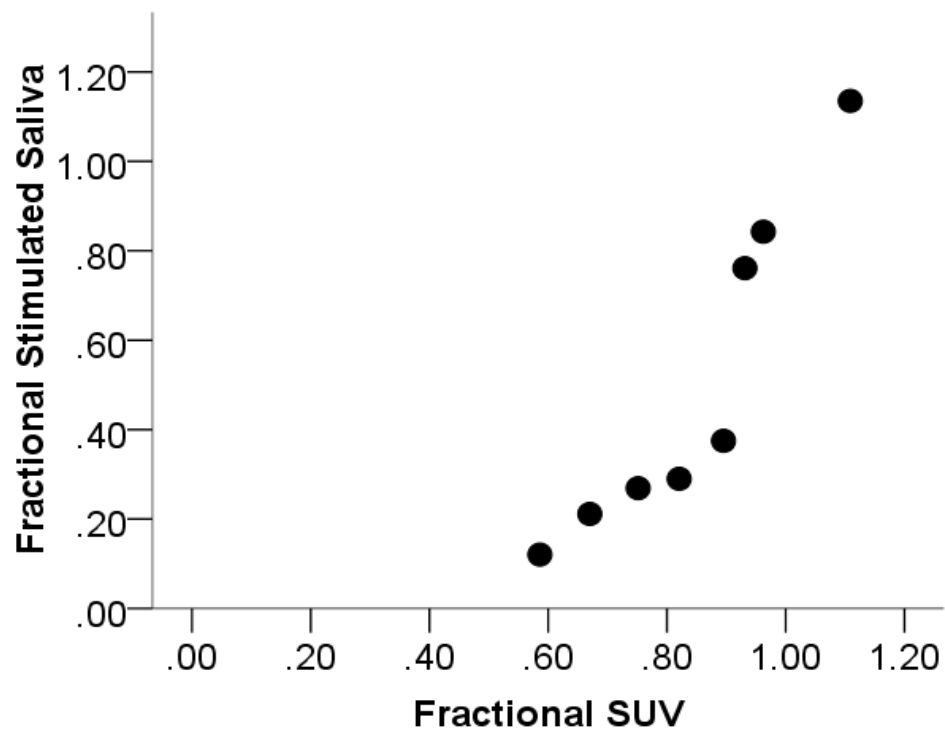


Figure 4.15 Fractional stimulated saliva mass and SUV

Fractional SUV was found to be positively correlated (Spearman's $\rho = 0.93$, $P < 0.01$) with fractional stimulated salivary mass.

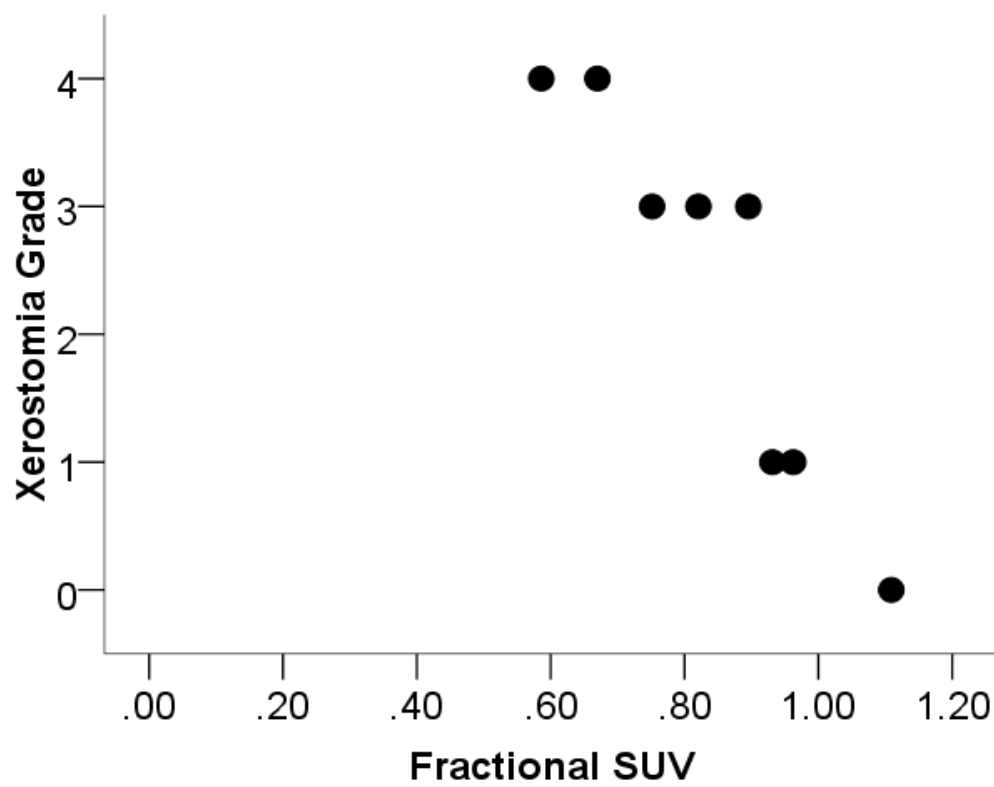


Figure 4.16 Fractional SUV and xerostomia grade

Fractional SUV was found to be negatively correlated (Spearman's $\rho = -0.964$, $P < 0.01$) to xerostomia grade as defined by the RTOG/EORTC late effects analytic scale.

Discussion:

Correlations between fractional parotid SUV and stimulated salivary mass suggest that Parotid SUV can function as an objective imaging biomarker of salivary function. Patients with parotid glands that had reduced SUV similarly had reduced stimulated salivary mass following RT. Further, parotid glands that exhibited increased SUV after RT paralleled increased stimulated salivary mass. Additionally, parotid gland fractional SUV had a strong negative correlation to xerostomia grade ($P < 0.01$). These data indicate that parotid SUV measured at the time of first follow-up after RT (approximately 8 – 9 weeks) may be a good surrogate for salivary toxicity.

As discussed in chapter 2, the parotid glands are composed of many spherical clusters of serous cells, known as the acini. The acinar cells are responsible for the production of saliva in the parotid glands. In studies of animal parotids, the glandular composition was altered greatly following irradiation. In one study of the rat parotid, the acinar cells comprised roughly 80% of the glandular volume prior to irradiation. Following the delivery of 30 Gy, the relative volume occupied by the acini was reduced to 20%. Reduction in the quantity of acinar cells, responsible for producing the watery secretions, corresponded to a 90% reduction in the saliva produced [71]. In humans, the loss of acinar cells at 10 – 12 weeks after the completion of RT has been shown to be the primary histopathological finding. For the current cohort, first follow-up salivary measurements were made seven weeks post-RT. While the salivary

mass measurement was earlier than the 10 – 12 week time frame cited above, a plausible explanation for the 59% reduction in stimulated salivary flow in 7/8 patients may be the reduction in the number of acinar cells. Furthermore, the acute phase of parotid damage characterized by the reduction in salivary flow and lack of alteration in glandular composition in the rat parotid was found to only last 0 – 10 days after irradiation.

In vivo and *in vitro* studies have demonstrated that uptake of FDG in tumors is directly related to the number of viable cells [99, 100]. Specifically, the numbers of cells expressing the GLUT-1 transporter correlate best with FDG uptake. GLUT-1 is a protein that facilitates transport of glucose across the cellular membrane. Expression of this protein is ubiquitous in human tissue and is responsible for basal glucose uptake. As the parotid glands are composed primarily of the serous acinar cells, which express GLUT-1, FDG uptake should be proportional to the number acinar cells. By consequence, parotid FDG uptake may be proportional to saliva production.

Although differences between parotid gland pre-RT SUV for patients deemed to have abnormal and normal baseline flow were statistically significant, this determination was arbitrary. Consider the patient classified as abnormal based upon baseline stimulated salivary mass. This patient produced a salivary mass of 2.2 g. This was based upon the calculated mass cutoff for a five minute saliva collection period. Based upon the low stimulated salivary mass, one might expect a correspondingly low parotid SUV. It was found that this patient

had the second lowest pre-RT SUV (1.0). On the other hand, the patient who produced the greatest baseline stimulated salivary mass (7.8 g) did not have the greatest baseline parotid SUV (1.6). More telling were the weak correlations between SUV and stimulated salivary mass. So while SUV may be related to the quantity of saliva produced, the lack of correlation between pre-RT SUV and stimulated salivary mass suggests that other confounding factors are most likely present. One possible factor is the imperfect correlation between SUV and FDG-uptake. Previous reports have shown correlation coefficients as low as 0.84 [99].

Other investigators have found absolute saliva mass is often not the variable that correlates closest with xerostomia [115]. Rather, it has been suggested that relative change in an individual's saliva production is more important. This logic seems reasonable in light of the wide range of baseline salivary masses that patients feel constitute normal salivary function (2.55 – 9.46 g). This was confirmed for the current cohort where the strongest correlations were found between relative changes in stimulated salivary flow and SUV (Pearson correlation = 0.93, $P < 0.01$). Finally, xerostomia grade was shown to be negatively correlated with fractional SUV (Spearman's $\rho = -0.964$, $P < 0.01$). This suggests that as patients have SUV values closer to baseline values, salivary toxicity should be reduced.

This study is limited in several respects. First, the imaging biomarker was validated against measurements of salivary flow, a metric that is known to have

large variability. However, little else exists in the way of objective measures of salivary toxicity. Second, the study was limited to a small number of patients. The protocol to establish a database of salivary toxicity from RT at The University of Texas MD Anderson Cancer Center is one of the first at this institution. With an accrual goal of 500 patients, the potential to fully investigate the association between SUV and saliva production is great. Finally, salivary output was measured as whole mouth salivary mass and not collected from individual glands. As detailed earlier, saliva results from the contributions of three major glands in addition to multiple minor glands [60]. Although fractional SUV correlated well with whole mouth stimulated salivary flow, saliva produced from other glands was possibly contaminating this measurement. However, this contamination was expected to be minimal. This is because the submandibular and sublingual glands receive large enough doses that saliva production after RT is effectively non-existent [42].

Conclusions:

In conclusion, pilot data suggest that parotid gland SUV would be well suited as an imaging biomarker for salivary toxicity. Relative change in parotid gland SUV correlated significantly with relative change in stimulated whole mouth salivary flow ($P<0.01$) and toxicity grade as defined by the RTOG/EORTC late effects analytic scale ($P<0.01$). The ability to assess parotid function using incidentally collected parotid uptake of FDG has great potential to decrease the reliance on poor measures of salivary toxicity like saliva collection and rarely used modalities like salivary gland scintigraphy. The results of specific aim 2a supported the hypothesis.

Chapter 5

Specific Aim 2b

To quantify and model the dose-metabolic response of the parotid glands in patients treated with IMRT for head and neck cancer using ^{18}F -FDG PET/CT

Working Hypothesis: Predicted post-RT parotid SUV will correlate with whole mouth stimulated salivary flow ($P < 0.05$).

Purpose

In the preceding study, Specific Aim 2a, the goal was to evaluate the potential use of SUV as an objective imaging biomarker of parotid function. In the small pilot cohort, it was observed that changes in SUV between baseline and follow-up correlated significantly with changes in stimulated salivary output. The application of novel imaging biomarkers holds great promise. However, no matter how well the biomarker may describe the physiologic process in question it is ultimately limited to describing the current functional status. In the case of Specific Aim 2a, the dose has been delivered and the parotid glands have been injured. Simply quantifying the damage can do little to alleviate the salivary toxicity.

The purpose of Specific Aim 2b, was to predict parotid post-RT SUV prior to the delivery of treatment. To accomplish this task, a phenomenological model of the dose-metabolic response relationship was created using only pre-RT metrics. The ability to estimate patient specific toxicity holds great promise for evaluation and refinement of radiation treatment plans.

Methods and Materials:

Patient Cohort

In order to construct a population based model of dose-SUV response, the patient cohort from Specific Aim 1 was exploited. As described earlier, these patients received pre- and post-RT PET/CT studies in order to assess and predict response to radiotherapy for head and neck cancers. The protocol did not specify the collection of salivary flow measurements, subjective salivary assessment through xerostomia quality of life questionnaires, or physician assessment of salivary toxicity. However, as elucidated earlier, the incidental collection of parotid FDG-uptake can have many tangential uses. Following this rationale, parotid SUV from patients enrolled in this protocol were utilized to model the dose-SUV response. Details of the cohort are provided in chapter 3.

To validate the dose-SUV response model, the patient cohort from Specific Aim 2a was employed. Patients enrolled in this IRB approved protocol had objective and subjective assessments of salivary function collected, as well

as, PET/CT imaging studies. Details of the validation cohort are provided in chapter 4.

Phenomenological Dose-Metabolic Response Model

The relative change in parotid FDG-uptake, expressed as fractional SUV, at 8 – 9 weeks after the completion of RT was modeled phenomenologically using pre-RT metrics. As shown previously, for the validation cohort, fractional SUV had strong and significant correlations to xerostomia grade and fractional stimulated saliva output. Modeling the continuous variable, fractional SUV, may allow a richer and more thorough evaluation of salivary toxicity than previous NTCP techniques that rely upon binary endpoints [116].

Two approaches were employed to model the dose-metabolic response. First, a mean dose-SUV model was assessed. Second, a voxel-based dose-SUV model was constructed. In both scenarios, fractional SUV was modeled as a function of a single parameter incorporating both planned physical dose and pre-RT SUV.

Metabolic Dose

To assess and model the functional dependence of fractional SUV, we introduced a figure of merit called metabolic dose. The metabolic dose was defined as,

$$D_{\text{Met}} = \text{Pre - RT SUV [Kg/L]} * \text{Physical Dose [cGy]}$$

where physical dose is the clinically planned dose distribution. The metabolic dose term weights the planned physical dose by the baseline SUV. In this manner, voxels with greater planned dose or initial metabolic function are expected to have greater reductions in baseline SUV than those voxels with lower planned dose or lower initial metabolic function. Assuming the parotid's density is equal to water, the units are in cGy.

The metabolic dose concept is illustrated in figure 5.1. In figure 5.1a, the relation of the parotid glands and the planned dose distribution is shown. From the image, it is clear that the most medial portions of the parotid glands will experience the highest doses. The most medial portions of the parotid glands are also observed to be the most metabolically active as evident in figure 5.1b. Consequently, the numeric value of the metabolic dose in this region will be the highest. As expected, the reduction of SUV in these regions was greatest as exemplified in figure 5.1d.

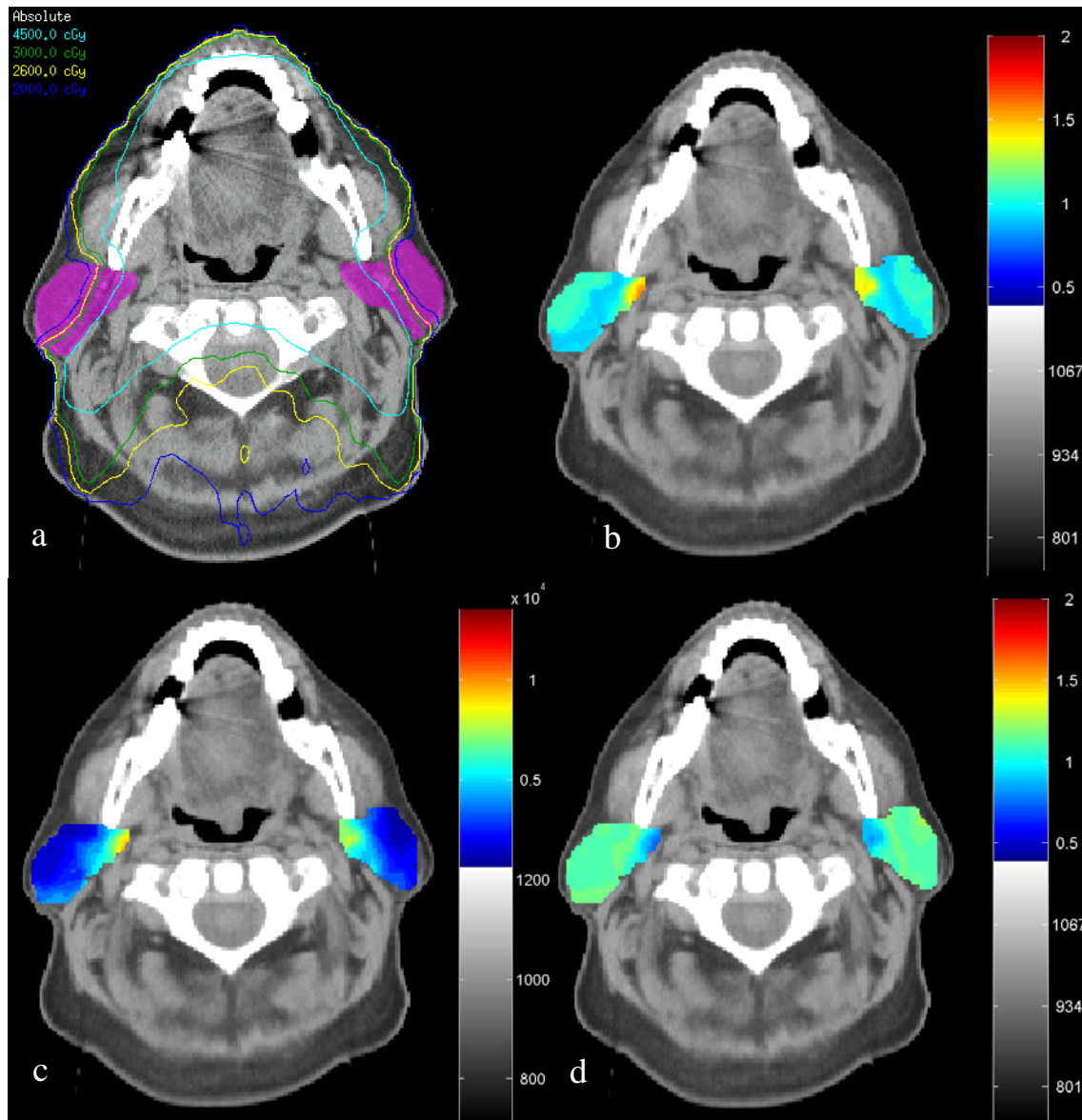


Figure 5.1 Metabolic Dose

Approved dose distribution with the parotids illustrated in purple. Skyblue denotes 45 Gy; green denotes 30 Gy; yellow denotes 26 Gy; and blue, 20 Gy. (b) Snapshot of baseline parotid SUV. The upper scale represents absolute pre-RT SUV. (c) Metabolic dose distribution for the parotid glands. (d) Biomarker distribution for the parotid glands after RT. In the upper scale, SUV is presented as a fraction of the baseline SUV. Of note is the correspondence between areas of highest DMet and lowest biomarker

Volume Delineation

As mentioned above, the parotid metabolic response model was constructed using the patient cohort from specific aim 1. The parotids for this cohort were contoured on planning CT images by the treating physician. For the mean dose-SUV model, the parotid contours were mapped to pre- and post-RT PET/CT images using the contour mapping workflow specified in specific aim 1.

In the case of the voxel-based model, the parotid contours created on the planning CT images were once again exploited. However, this model required one-to-one correspondence between pre- and post-RT SUV as well as dose at the voxel level. To facilitate this, the enhanced demons algorithm was utilized, but rather than mapping contours the image sets were deformed. Briefly, the deformable transformations that resulted from CT-to-CT alignment were applied to the corresponding PET images. As a result, the pre- and post-RT PET images were aligned to the planning CT image. The detailed methodology to perform this within the CAT workspace is discussed in the next section.

Co-registration of Planned Dose and SUV

The following methodology details the steps required to align PET/CT images to a reference image set within the CAT image registration workspace. While not a fusion image viewer, the CAT workspace can be utilized to collect SUV (Body weight, lean body mass, or body surface area normalization) or uptake for a volume of interest.

All serial PET/CT studies were identified in Stentor and uploaded to Evercore for retrieval as described earlier. Prior to importing original patient treatment plans into the CAT workspace, it was necessary to contour the second cervical vertebrae for image alignment purposes. Contouring was performed using Philip's Pinnacle³ treatment planning software, version 7.6c. After opening the CAT workspace, the original patient treatment plans were imported as detailed earlier.

The images were then preprocessed. All serial PET/CT studies were first checked for concordance of slice spacing, thickness, and pixel size. In the event that one of these imaging parameters did not match, the Image Guided Radiation Therapy (IGRT) Utilities software package created by our group was utilized to re-sample the images appropriately. Next, the planning CT imaging parameters were inspected for concordance with the PET/CT parameters. Again, if the parameters differed, the image was re-sampled appropriately. By matching the imaging parameters of each data set we can achieve voxel-by-voxel correspondence for the planning CT and serial PET/CT studies.

The next steps resemble those described in chapter 3 for the contour deformation methodology. First, the bony alignment algorithm discussed in chapter 3 was utilized to align the C2 vertebrae of the reference and daily image sets. To deform the PET/CT study of interest, the advanced options were used from the deformation dialog box and batch registration was selected. Most importantly, the direction of deformation was specified as Daily -> Reference.

The DICOM PET images corresponding to the current PETCT study were then selected. The deformation field mapping the Daily CT to the Reference CT was applied to each voxel in the PET image, producing a PET image aligned to the planning CT.

Before dose, pre- and post-RT SUV can be extracted from each contour the dose from the approved treatment plan must be collected. Using the *IGRTDumpDose* script in the Pinnacle treatment planning system extracted the required data. Finally, using the advanced options menu, the “Dump PET & Dose” function was utilized to extract data for every voxel within the contours contained in the alignment set. The data were output to a text file containing each voxel location along with dose, pre- and post-RT SUV.

Validation of the Models

After a mathematical form of the dose-metabolic response relationship was selected, the model was validated using the patient cohort from specific aim 2a. Initial validation consisted of evaluation of the R^2 statistic. However, graphical techniques are often the most telling validation of a models performance. For both the mean and voxel-based dose-SUV response models, plots of predicted versus observed fractional SUV were constructed. For a perfectly performing model, data points on these plots should lie along the line $y = x$. Therefore, a well performing model should have a strong linear correlation

between the predicted and observed values. Spearman's correlation was utilized to assess these correlations and statistically significant results were defined as having $P < 0.05$. Relative errors were computed as, $(\text{Predicted} - \text{Observed})/\text{Observed} * 100$.

Results:

Fractional SUV and Metabolic Dose

For the mean model, the mean planned dose and pre-RT SUV of the parotid glands was used. In figure 5.2, the fractional SUV is plotted as a function of these two variables. The average dose delivered to the parotid glands for the population was 25.4 ± 6.0 Gy. Mean dose to the parotid glands ranged from 14.4 – 42.8 Gy. The clustering of the mean parotid doses is an artifact of the planning goal to keep the mean dose less than 26 Gy. The fractional SUV observed for the population had an average value of 0.96 ± 0.24 . The fractional SUV range stretched from a minimum value of (0.62) to a maximum of (1.85). Overall, these data indicated that FDG-uptake was reduced following delivery of treatment. However, a simple relationship between dose and fractional SUV did not seem to be present.

The relationship between fractional SUV and the metabolic dose figure of merit is shown in figure 5.3 for the population. The overall shape of the curve suggests that fractional SUV has a consistent relationship to the figure of merit and matches well with the expected dependence. As stated above, larger

metabolic dose values were expected to have correspondingly greater reductions in SUV. This would translate into smaller fractional SUV values. Examining the extremes of the plot, a D_{Met} value for one patient of interest was calculated to be (8667) with a fractional SUV of (0.73). Whereas an example patient from the lower extreme with a D_{Met} value of (2901) had a fractional SUV value of (0.90).

The raw data from figure 5.3 were binned for the purpose of modeling. Five bins were utilized to capture the overall shape of the curve. The width of each D_{Met} bin was 2600. In figure 5.4, the mean value of each bin with one standard deviation is displayed.

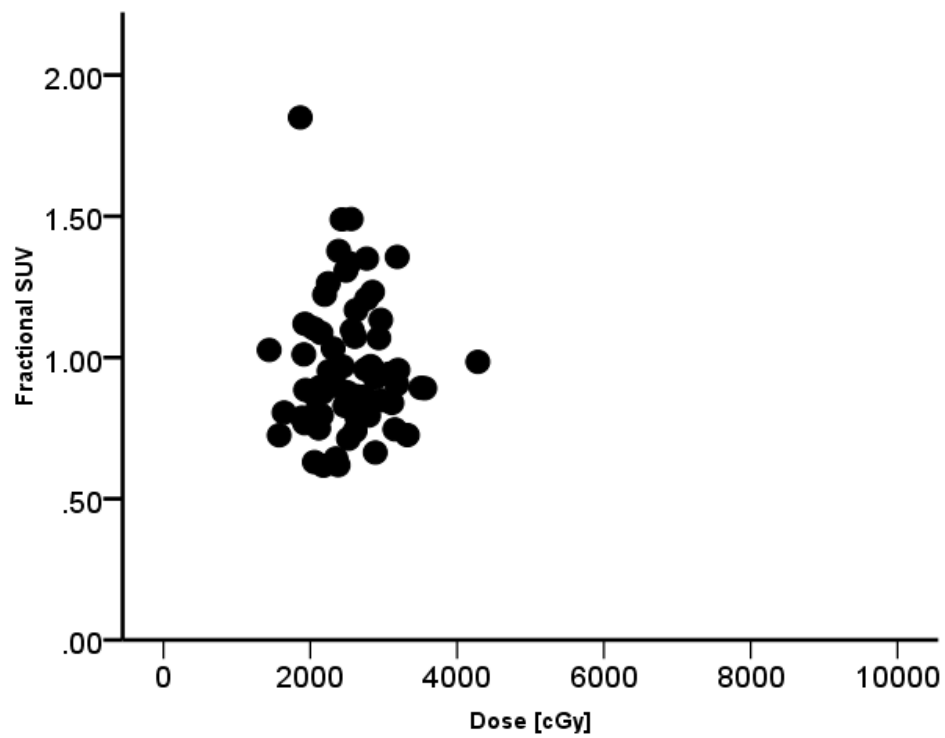


Figure 5.2 Fractional SUV dose response

Fractional SUV is plotted as a function of whole organ planned mean dose.

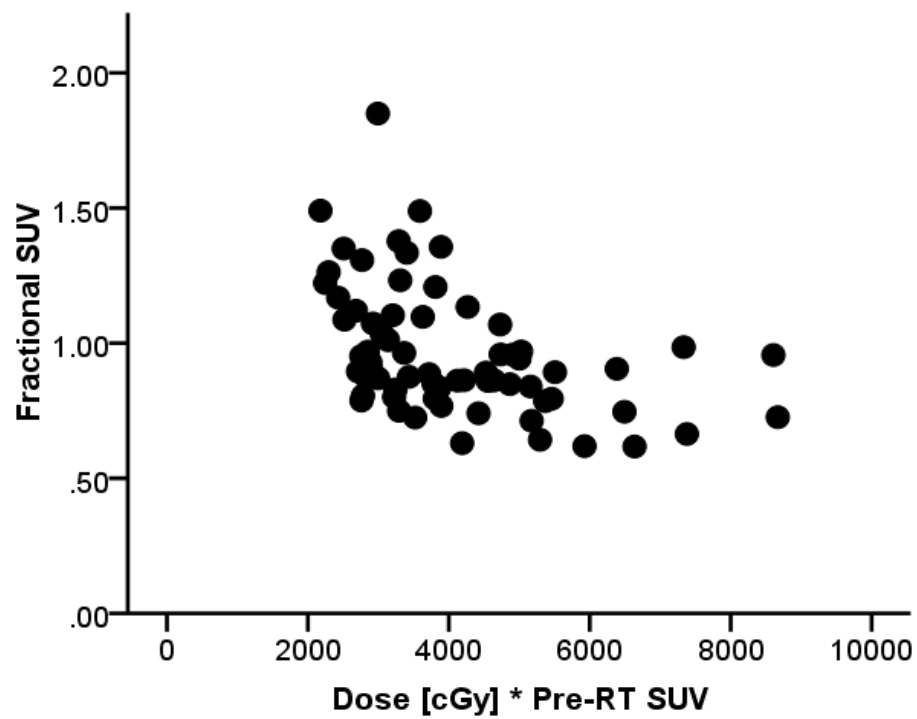


Figure 5.3 Fractional SUV and D_{Met}

Fractional SUV is plotted as a function of the metabolic dose figure of merit. $DMet$ is shown deconstructed in this plot.

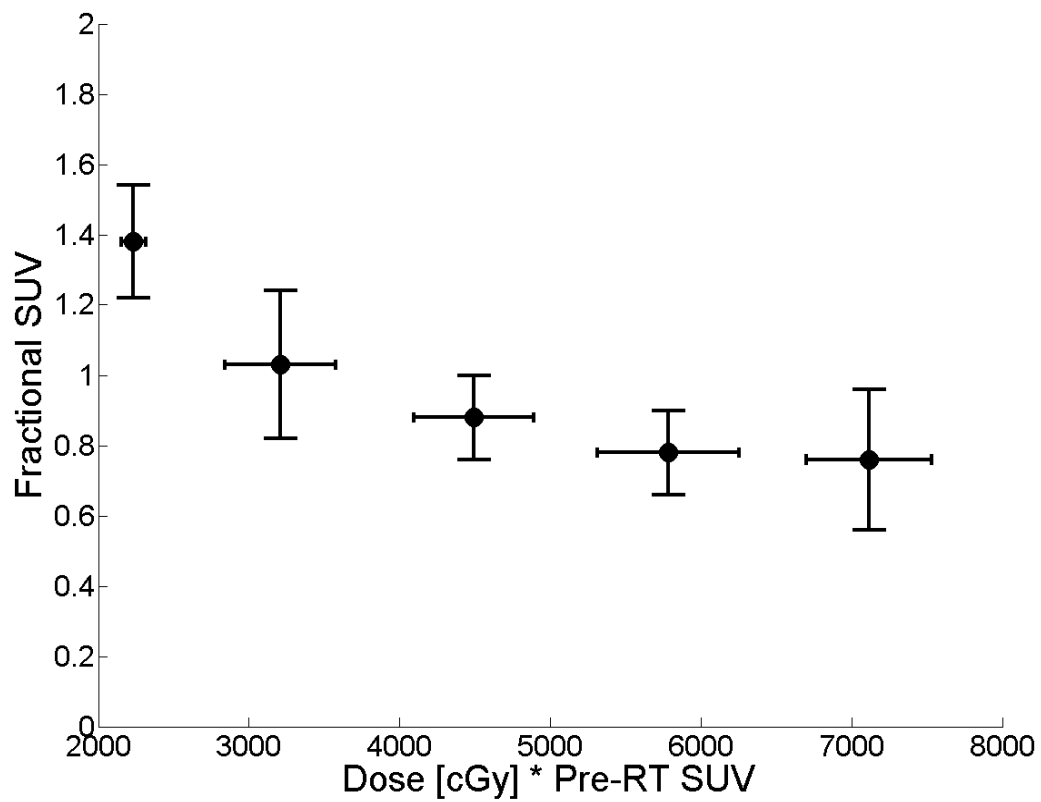


Figure 5.4 Fractional SUV and D_{Met}

The average fractional SUV is plotted against the mean metabolic dose for each bin of width 2600. The error bars represent one standard deviation.

Validation of the Mean Model

For the mean dose-SUV model, the fractional SUV was modeled over a range of $D_{\text{Metabolic}}$ values extending from a minimum of 2235 to a maximum of 7115. A power law with a constant offset was selected to model the data over the range of values found in this population. The functional form of this model is

$$f(x) = a * x^b + c$$

where a , b , and c represent the coefficients. In the above equation, x represents the metabolic dose figure of merit. The results of the non-linear least squares fit are shown in figure 5.5. The constant a value was calculated to be 3.6×10^5 with 95% confidence bounds of -2.8×10^6 , 3.6×10^6 . The constant b value was found to have a value of -1.7 with 95% confidence bounds of (-2.88, -0.52). Finally, the offset constant, c , had a value of 0.65 with 95% confidence bounds of (0.42, 0.87). The adjusted R^2 statistic for the fit was found to be 0.996. The sum of squares due to error (SSE) was found to be 0.0004.

Next, the patient cohort from specific aim 2b was used to further validate the phenomenological model. In figure 5.6, a scatter plot of the observed fractional SUV for this cohort versus the fractional SUV predicted from the mean

dose-SUV model is shown. The dotted line represents a perfectly performing model. The mean relative error between the predicted and observed fractional SUV was 13%. The maximum and minimum relative errors found were 76% and -13%, respectively. The most accurately predicted fractional SUV had a relative error of 0.07%. For the patient with the largest relative error the observed fractional SUV was (0.47), but the predicted value was (0.83). However, this patient was classified as having “abnormal” baseline salivary function. Finally, the predicted and observed values were found to have a strong and significant correlation, Spearman’s $\rho=0.71$, $P<0.01$. The combination of fit statistics and graphical analysis indicate the phenomenological model performs well.

Lastly, the correlation between the predicted fractional SUV and the observed fractional stimulated salivary output was assessed. In figure 5.7, the fractional stimulated salivary mass is shown to be significantly correlated with predicted fractional SUV, Spearman’s $\rho = 0.79$, $P<0.05$. As observed in specific aim 2a, only one patient had SUV that increased beyond baseline and this corresponded to an increase in stimulated salivary mass. However, the mean model predicts that two patients would have SUV increases from baseline.

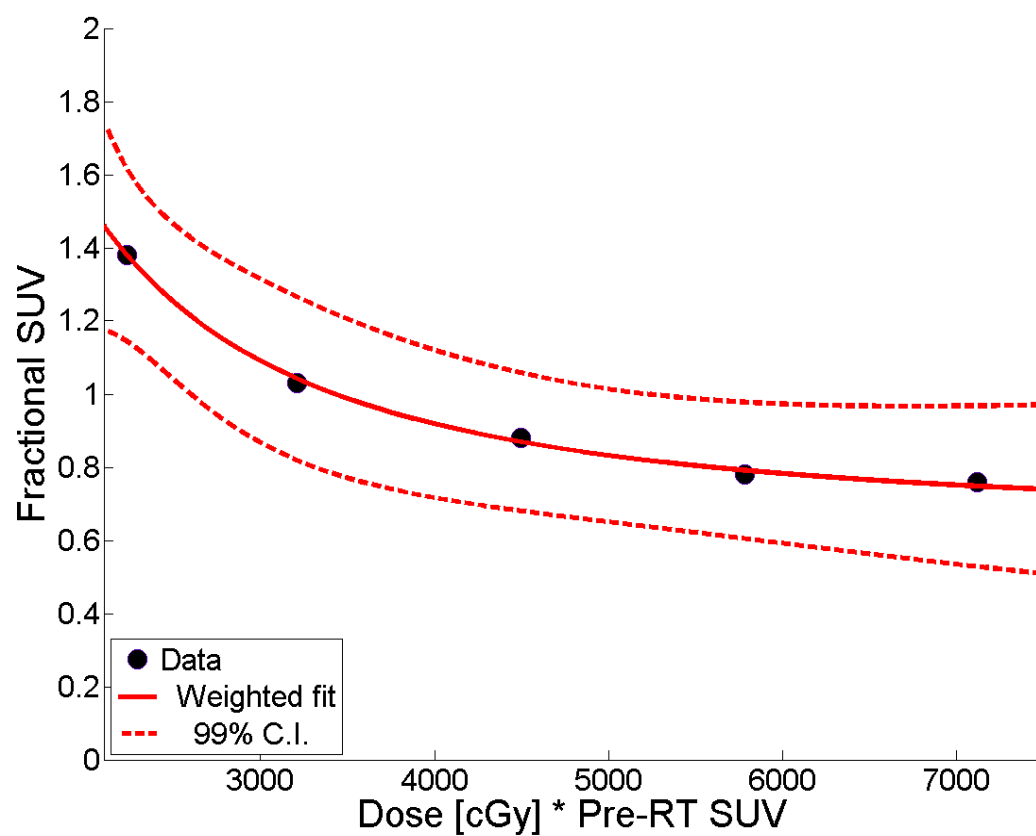


Figure 5.5 Fitting results for the mean model

Mean metabolic dose and fraction SUV data are plotted with the resulting weighed fit for a constant offset power law.

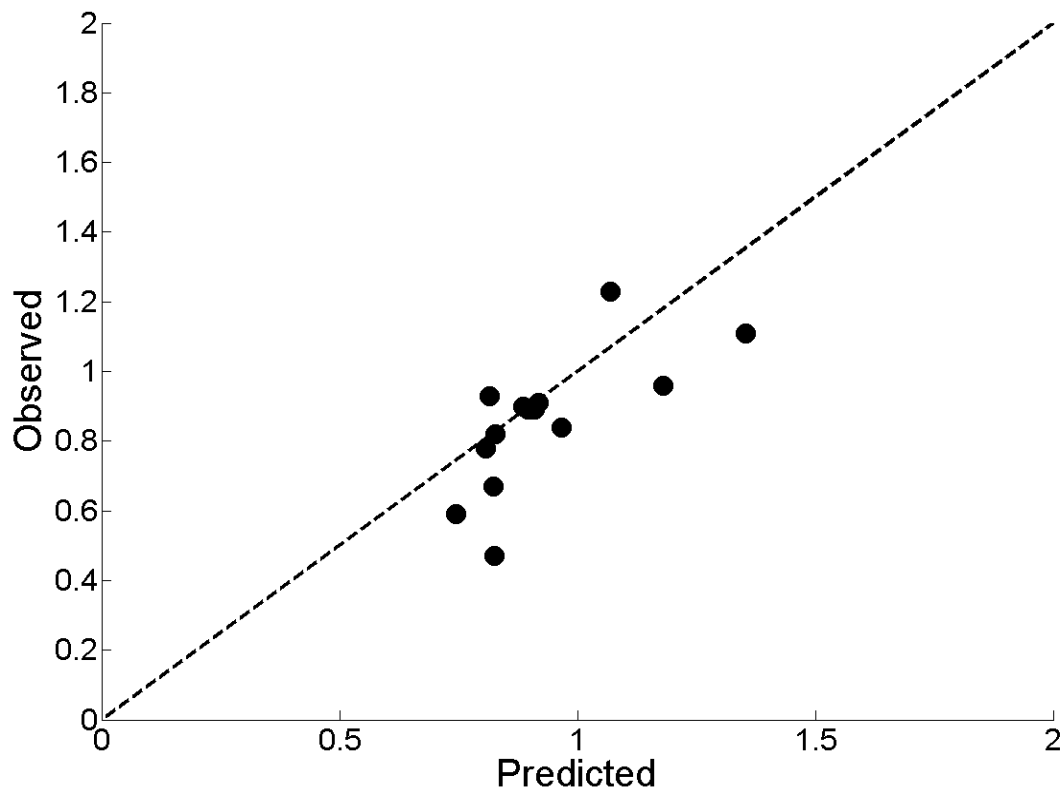


Figure 5.6 Mean model validation

The observed fractional SUV values are plotted against the fractional SUV values predicted from the mean model (Spearman's $\rho = 0.71$, $P < 0.01$).

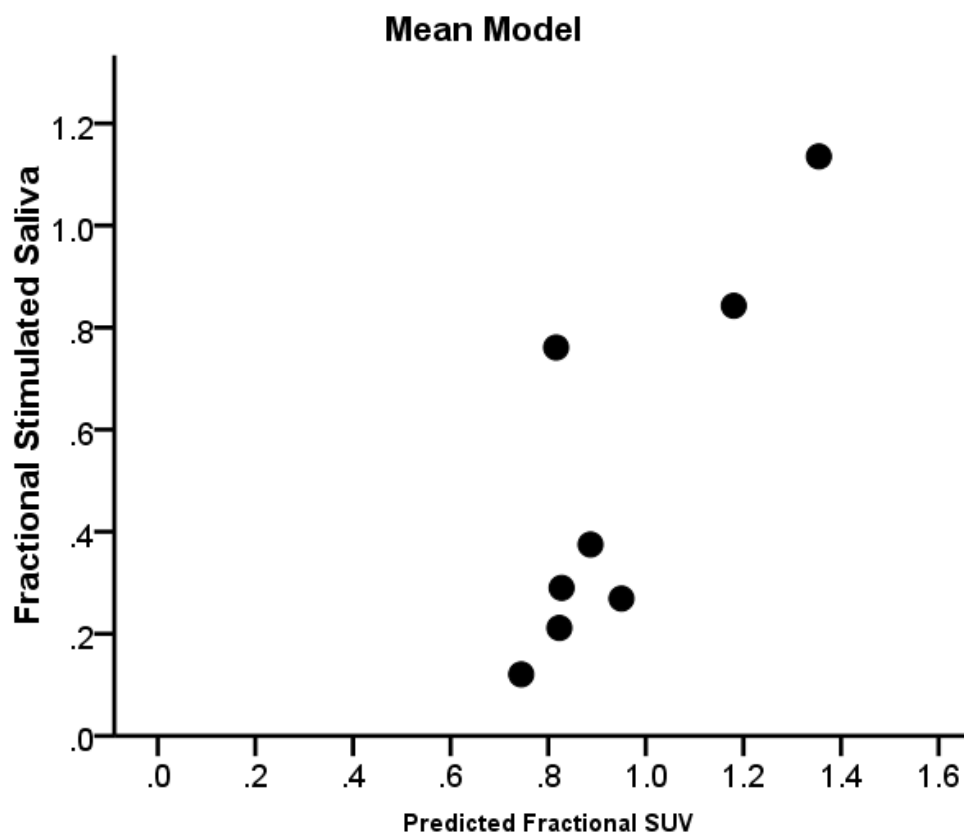


Figure 5.7 Predicted SUV and observed saliva changes

The observed fractional stimulated saliva mass was plotted as a function of the predicted fractional SUV (Spearman's $\rho = 0.79$, $P < 0.05$).

Voxel-based Fractional SUV and Metabolic Dose

The second technique used to construct the phenomenological model was voxel-based. After co-registration of pre- and post-RT SUV with planned dose, the voxel data were binned. The bin width was set at 2000, but expanded for the highest metabolic dose values to increase the number of voxels in the most starved bins. In figure 5.8, the relationship between the metabolic dose figure of merit and fractional SUV as determined on a voxel-by-voxel basis may be examined. The error bars represent one standard deviation.

Once again a consistent relationship between the two variables was observed. Furthermore, the shape of the distributions also conformed to the initial expectations of an inverse relationship between the fractional SUV and the metabolic dose figure of merit. Of particular interest is the much wider range of metabolic dose values observed. For the mean model, the metabolic dose metric range was up to 6.5×10^3 , whereas the range for the voxel model was an order of magnitude larger, up to approximately 7.8×10^4 . As alluded to earlier, the narrow range of mean doses delivered to the parotid glands is the primary reason for the constricted range. However, on a voxel-by-voxel basis the dose range was much larger, 67 Gy versus 28 Gy found for mean planned doses to the whole parotid glands. Because of this, the voxel model is able to account for dose heterogeneities within the parotid glands.

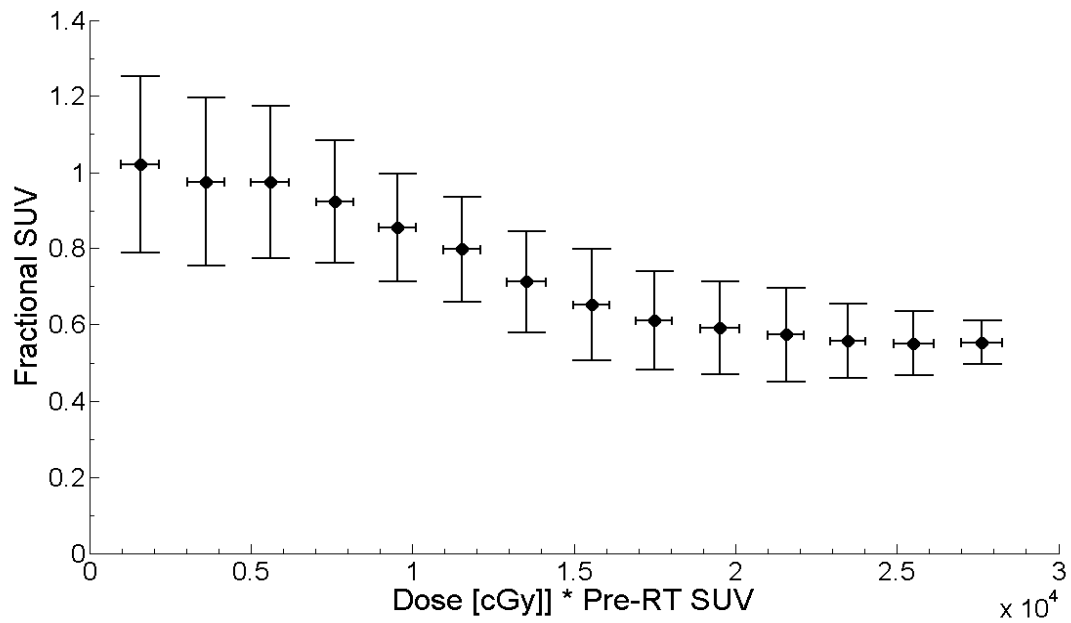


Figure 5.8 Voxel-based dose-SUV response

The fractional SUV-metabolic dose relationship for the entire cohort was compiled for the entire cohort.

Validation of the Voxel Model

The A log transform of the fractional SUV and metabolic dose was utilized to aide in the fitting. A modified logistic function was utilized to model fractional SUV at the voxel level as a function of metabolic dose. The functional form has been modified to prevent fractional SUV values of zero. Additionally, a boundary condition was imposed on the voxel data. At zero metabolic dose the fractional SUV is assumed to be one. This means that if the patient were to receive no dose the SUV value should remain at baseline values. The functional form of this model is

$$f(x) = \frac{a}{1 + \left(b * \left(\frac{x}{c} \right)^d \right)} + e$$

where, **a-e** represent the coefficients. In the above equation, x represents the metabolic dose figure of merit. The results of the linear least squares fit are shown in figure 5.9. The constant **a** value was calculated to be 0.4669 with 95% confidence bounds of (0.4434, 0.4905). The constant **b** value was found to have a value of 7.599 with 95% confidence bounds of (5.287, 9.912). The constant **c** that scales the metabolic dose value was found to have a value of 2.25×10^4 with 95% confidence bounds of $(2.1 \times 10^4, 2.8 \times 10^4)$. The power value,

constant **d**, was found to have a value of 3.943 with 95% confidence bounds of (3.181, 4.706). Finally, the offset constant, **e**, had a value of 0.5324 with 95% confidence bounds of (0.513, 0.5517). The adjusted R^2 statistic for the fit was found to be 0.9987. The sum of squares due to error (SSE) was found to be 0.0007.

Next, the patient cohort from specific aim 2b was again used to validate phenomenological model. In figure 5.10, a scatter plot of the observed fractional SUV for this cohort versus the fractional SUV predicted from the voxel-based dose-SUV model is shown. The dotted line represents a perfectly performing model. The mean relative error between the predicted and observed fractional SUV was 9.6%. The maximum and minimum relative errors found were 65% and -22%, respectively. The most accurately predicted fractional SUV had a relative error of approximately 0.2%. The patient with “abnormal” baseline salivary function also plagued the voxel-based model. For this patient the observed fractional SUV was (0.47), but the predicted value was (0.65). The voxel-based model was able to improve the predicted fractional SUV for this patient, but still failed to fully describe the extent of SUV reduction. There were two addition patients where fractional SUV was noticeably overestimated, 46% and 25%. The fractional SUV was underestimated markedly for one patient. The observed fractional SUV in this case was (1.23), but the predicted value was only (0.96). This represented a roughly 22% underestimation. This patient also had “abnormal” baseline salivary function. Finally, the predicted and observed values were found to have a strong and significant correlation,

Spearman's $\rho=0.94$, $P<0.01$. The combination of the adjusted R^2 and graphical analysis indicate the voxel-based phenomenological model performed at least as well as the mean dose-SUV model and in many cases better. Overall, the relative error was reduced from 13% to 9%.

Lastly, the correlation between the predicted fractional SUV and the observed fractional stimulated salivary output was assessed. In figure 5.11, the fractional stimulated salivary mass is shown to have a strong and significant correlation with our figure of merit, Spearman's $\rho = 0.85$, $P<0.01$.

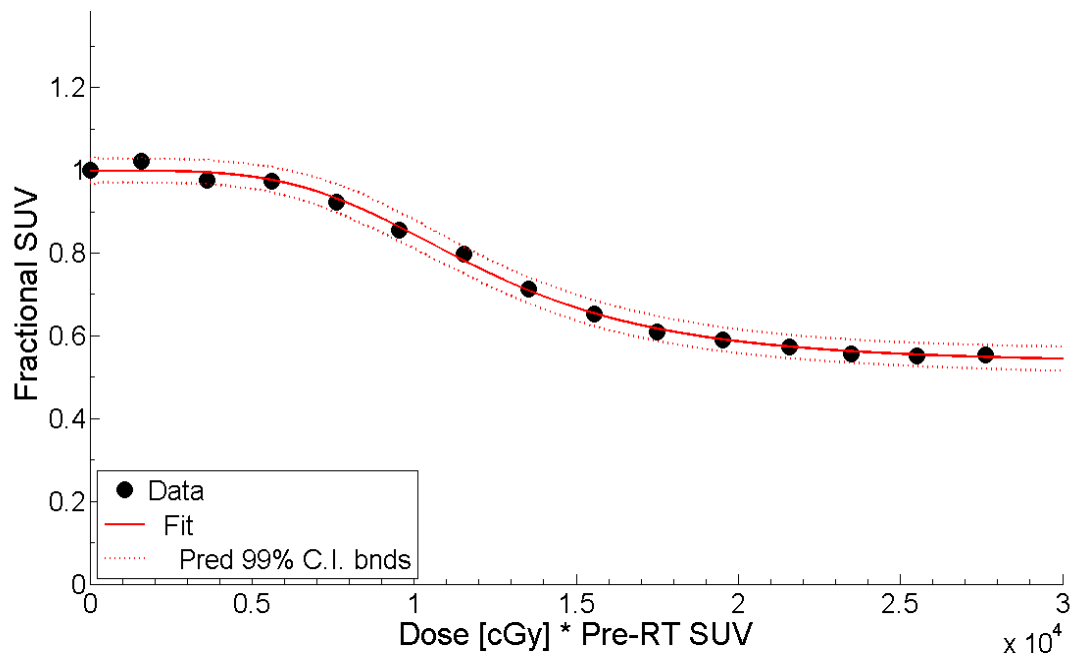


Figure 5.9 Fitting results for the voxel model

Mean metabolic dose and fraction SUV data are plotted with the resulting modified logistic model fit.

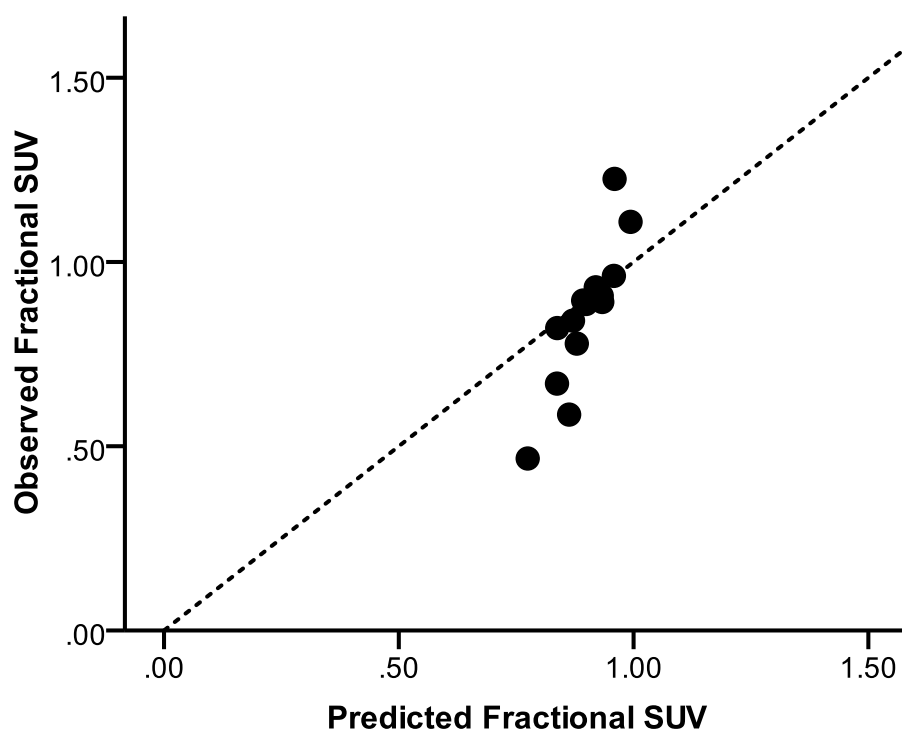


Figure 5.10 Voxel model validation

The observed fractional SUV values are plotted against the fractional SUV values predicted from the mean model (Spearman's $\rho = 0.94$, $P < 0.01$).

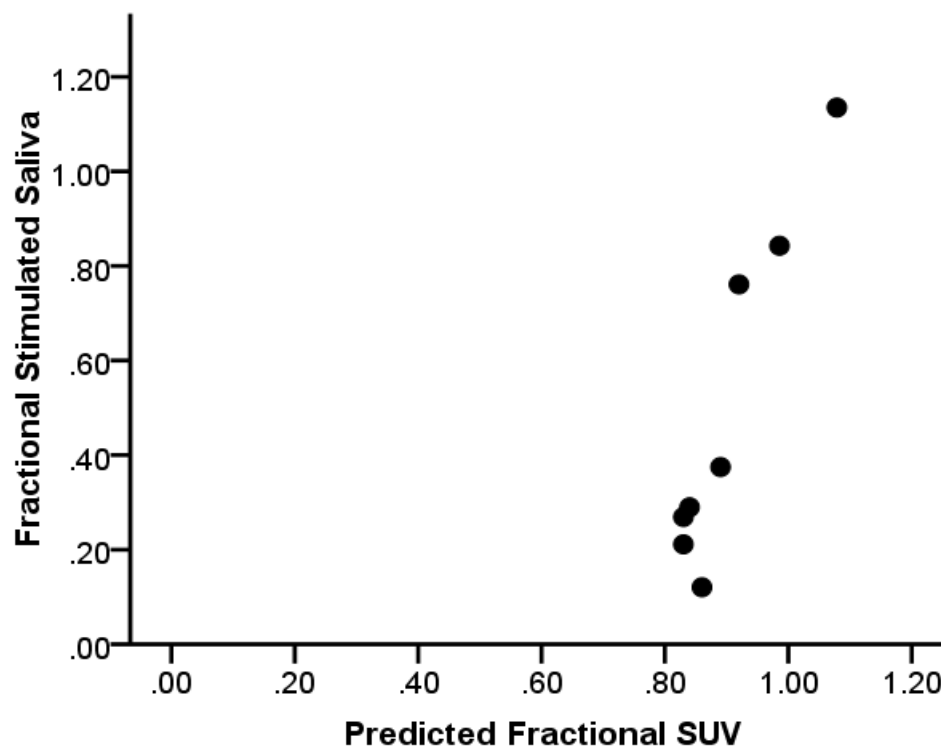


Figure 5.11 Predicted SUV and observed saliva changes

The observed fractional stimulated saliva mass was plotted as a function of the predicted fractional SUV (Spearman's $\rho = 0.85$, $P < 0.01$).

Discussion

Predicting normal tissue toxicity has long been a goal to improve patient outcomes [117]. Fully understanding the response of normal tissues is vitally important in determining tolerance doses. Dose response modeling utilizing *in vivo* biomarkers offers the possibility for patient specific estimations of toxicity based upon biologic function. Additionally, in comparison to other image-based dose-response studies, PET/CT offers a 3D technique that many clinicians have routine experience with [4, 45]. However, exploiting ^{18}F -FDG PET/CT for dose-response modeling is still in an embryonic stage [106, 107].

The first model developed as part of this thesis work was a mean parotid dose and SUV model. Although the range of mean doses delivered to the parotid glands only spanned 28.4 Gy, mean dose to the parotid glands has been shown to be one the strongest predictors of salivary toxicity [39]. However, as shown in figure 5.2, a simple relationship between Dose and relative SUV change in the parotid glands was not evident. We hypothesized that in addition to the mean dose delivered to the parotid glands, the baseline function of the glands should also be relevant. A figure of merit called the metabolic dose was introduced to capture this interplay between the spatial dose distribution and glucose metabolism. Because SUV is known to be proportional to the number of viable cells, dose delivered to more metabolically active regions were weighed more heavily with this metric and expected to experience greater reductions in SUV after treatment. Although the selection of the metabolic dose metric was

serendipitous, motivation for its selection can be found through multivariate fractional polynomial regression analysis [118, 119]. This type of regression analysis is useful when the covariates in the model may be non-linear. Utilizing the statistical package R, the influence of dose, pre-RT SUV, and the metabolic dose metric on regression models of fractional SUV was assessed using the voxel-based data. While none of the three covariates listed were eliminated from the final model, the coefficients were illuminating as to the best metric for SUV response modeling. For dose, the coefficient was -6.7×10^{-8} and the coefficient for pre-RT SUV was 8.5×10^{-2} . However, when the product of dose and pre-RT SUV, metabolic dose, was considered the coefficient was -2.5. The coefficient for the metabolic dose metric was at least two orders of magnitude larger than the separated covariates. This indicates that metabolic dose is the dominate covariate and supports the use of this figure of merit for modeling the SUV response relationship. A plot of metabolic dose and parotid SUV relative change displayed a consistent relationship that conformed to expectations. The mean model was fit with a constant offset power law. The fit statistics indicated a good fit, but because the metabolic dose metric and fractional SUV are correlated this was expected.

One of the potential drawbacks of the phenomenological model is that it was only applicable over the range of observed data. Beyond these limits the model is ill behaved. For instance, as the metabolic dose increases, the fractional SUV will asymptotically approaches the constant offset, which for this cohort was 0.65. However, suggesting that parotid SUV relative change has a

minimum of 65% of baseline may be unrealistic. Interesting and perhaps unsurprisingly, given the fractional SUV dependence observed, only three parotid glands breached this barrier in the model cohort. For the test cohort, two patients were below the minimum value of 65%, one of which had abnormal baseline salivary function. At the other extreme, as metabolic dose values approach zero the relative change goes to infinity. This seems to be physiologically unlikely. Increases in glucose metabolism are possible and were observed, but setting a ceiling value for increase is unclear. In one study by Schwartz-Arad et al., the effects of compensatory hyperplasia following partial submanibulectomy in rats found increased salivary function and an approximately 154% increase in the number of acinar cells the remaining submandibular gland [120].

Graphical validation of the mean model was favorable. Scatter plots of the predicted and observed fractional SUV for the mean model demonstrated highly positive correlation ($\rho = 0.71$, $P < 0.01$). Furthermore, because the ultimate goal is to predict the fractional SUV that would result from treatment prior to its actual delivery, observed changes in stimulated salivary mass were plotted against predicted changes in SUV. Although highly correlated (Spearman's $\rho = 0.79$, $P < 0.05$), when compared to figure 4.15 in chapter 4, it is apparent that the relation between fractional SUV, which serves as the imaging biomarker, and relative change in stimulated salivary mass was altered. Specifically, when fractional SUV was evaluated as a biomarker only one individual had an increase in SUV and this corresponded to an increase in stimulated salivary

mass. For the predicted values, two patients were predicted to have increases in SUV and it would be expected that both might have increased salivary output. However, this was not true for one of these patients and the actual stimulated salivary mass decreased by almost 20%. Overall the mean dose-SUV response model represented an easily implementable method that, with further validation, may allow for individualized assessment of the salivary toxicity resulting from a treatment plan design.

The second method used to construct the dose-SUV response model was based upon voxel-by-voxel correspondence of dose and SUV. The benefit of this model is the ability to account for subtle differences in the spatial distribution of planned dose. Figure 5.12 illustrates this with two similar dose volume histograms that led to very different functional outcomes. The range of doses available at the voxel level was much larger than those available when whole organ mean dose was used (67 Gy vs. 28 Gy). Because of this, a richer picture of the metabolic response was possible. This was reflected in the reduced relative error for the voxel-based model (9% vs. 13%). An examination of the observed vs. predicted fractional SUV plot showed that the model performed quite well (Spearman's $\rho = 0.94$, $P < 0.01$).

The voxel-based relationship between the metabolic dose figure and fractional SUV suggested a sigmoidal dependence. The general trend of this data matches with that found in scintigraphic studies of parotid gland excretion as measured by reductions in tracer uptake. The voxel-based model differed

from the mean model in that whole organ function was reconstructed from sub-glandular response. By predicating sub-glandular response, the voxel-model was able to account for variations in the spatial distribution of dose. However, this also meant that the voxel-based model was more susceptible to the accuracy of the dose calculation engine. One recent study compared the Pinnacle³ version 7.6c superposition/convolution algorithm with Monte Carlo (EGS4) calculations for head and neck IMRT patients. Mean doses to the ipsilateral and contralateral parotid only differed by -0.5% and -0.04%, respectively [121]. For the mean dose model, errors in the calculated dose were minor. However, calculated doses from the surface down to 2 mm depth were shown to have relative errors that exceeded 5%. The most superficial border of parotid glands was found at approximately 2 mm depth. Even with these potential differences between the delivered and calculated dose, whole organ fractional SUV constructed from sub-glandular response agreed well with observed values of SUV.

Additionally, the ability of the voxel-based model to account for potential hyperplasia may be limited. As seen in figure 5.8, the voxel data failed to show evidence of SUV increases from baseline for low metabolic dose values. However, parotid mean SUV was observed to increase from baseline. If the voxel-based model is utilized to reconstruct parotid mean SUV it will fail to account for observed increases.

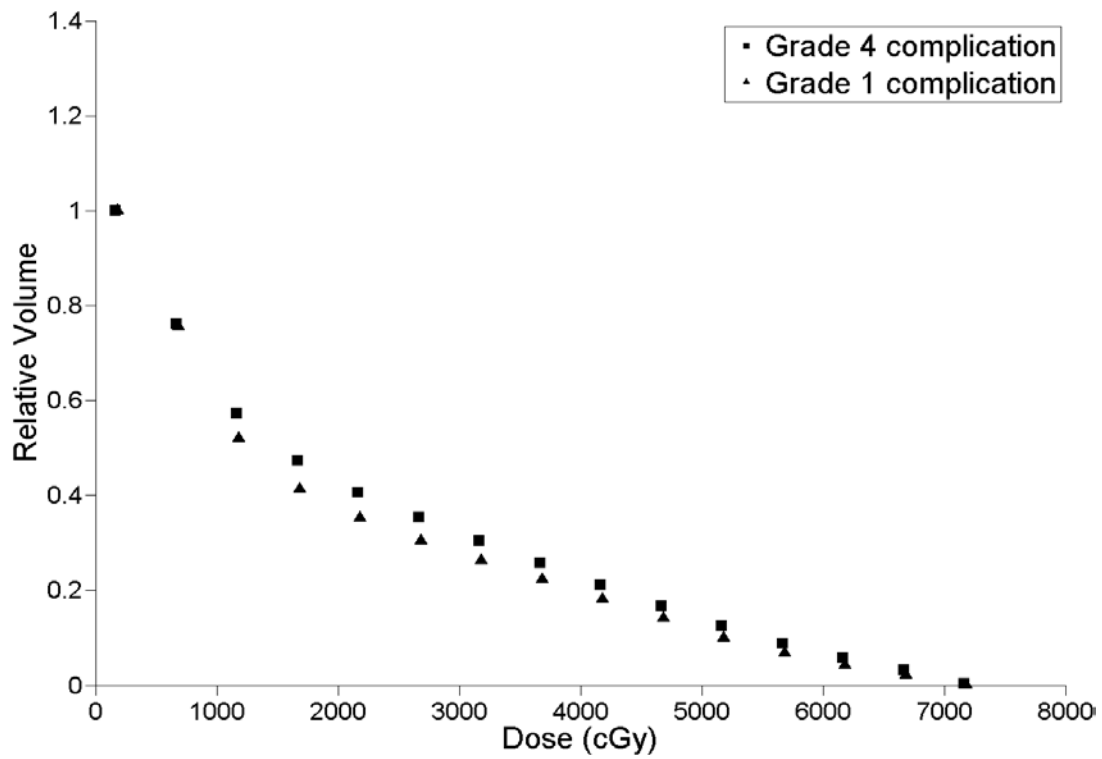


Figure 5.12 Parotid DVH comparisons

For a high (square) and low (triangle) oral complication patient, the cumulative DVH displays surprisingly subtle differences. Both patients had a whole organ planned mean dose less than 26 Gy.

Consider a patient that had an observed increase in SUV of 23%. The mean model predicted an increase of 10% while whole organ SUV reconstructed from the voxel-based model predicted an SUV decrease of 4%. For the mean dose and SUV data, increases in SUV from baseline were observed.

Conclusions

In conclusion, the dose-SUV response relationship in the human parotid glands was quantified and modeled. Utilizing only baseline metabolic function and the planned dose distribution, predicting parotid SUV change or salivary toxicity, based upon specific aim 2, became possible. Two techniques were used to create the phenomenological models: 1) based upon whole organ planned mean dose and pre-RT SUV and 2) based upon dose and SUV at the voxel level. Mean and voxel-based models performed well at predicting SUV changes in a validation cohort with relative errors of 13% and 9%, respectively. Predicted SUV changes and observed stimulated saliva mass changes correlated strongly for both models although some shifting of the relationship observed in specific aim 2a was observed. The results of this aim supported the hypothesis.

Summary and Conclusions

Central Hypothesis:

We hypothesize that more accurate image registration will lead to improved quantitation of tissue response to radiotherapy with ^{18}F -FDG PET/CT imaging.

Conclusions

The central hypothesis of this work was an overarching statement that connected the myriad applications of ^{18}F -FDG PET/CT investigated in this work. As described in the pages above, quantitative treatment response monitoring with ^{18}F -FDG PET/CT holds great promise. Colloquially, treatment response monitoring has referred to analysis of only neoplastic tissue. However, such a narrow application of a paradigm changing technology will limit novel ways to improve patient management. As such, this work was designed to improve upon the existing uses of this modality and expand its use to new and interesting vistas.

In the first specific aim, the purpose was to apply deformable image registration techniques to a study designed to assess and predict response to

RT using PET/CT. Total tumor volumes, defined on CT and PET images, and the most metabolically active volumes were investigated to find the most accurate method to judge response. Secondly, two methods of aligning these volumes to serial PET/CT studies were tested. The first relied upon a deformable image registration algorithm, while the second employed a more conventional bony alignment technique. The deformably mapped SUV_{Peak} volume was the best metric for predicting response to RT. The PPV (63%) of this metric was greater than post-RT contrast enhanced CT (50%) and all other PET metrics (40%, 21%, and 57%). When compared to bony alignment of the SUV_{Peak} volume, deformable image registration led to significantly improved prediction of response to RT. Specific aim 1 supported the working hypothesis that voxel-by-voxel mapping of primary volumes to serial F-FDG PET/CT images would improve region based tumor response analysis.

In the second specific aim, the goal was to evaluate the use of a novel imaging biomarker for salivary function following RT. To our knowledge, this was the first time the relationship between salivary output and salivary gland uptake of ¹⁸F-FDG has been investigated. Relative change in parotid gland SUV was found to have strong and significant correlations to relative change in stimulated salivary mass (Pearson correlation = 0.93, P<0.01). Utilizing the RTOG/EORTC Late Effects analytic scale, patients were stratified into xerostomia grade. Although limited by patient number, individuals with parotid SUV closer to baseline values were found to have lower grade xerostomia at the time of first follow-up (Spearman's ρ = -0.96, P<0.01). Mapping pre-RT parotid

volumes across serial imaging studies yielded a novel method to assess salivary toxicity due to RT. Specific aim 2a supported the hypothesis that parotid gland SUV would function as an objective imaging biomarker of salivary toxicity, correlating with stimulated salivary flow ($P < 0.05$).

In the third specific aim, the goal was to construct a model of the dose-SUV response relationship in the parotid glands to allow for pre-treatment prediction of post-treatment parotid SUV values. Utilizing a figure of merit that captured both the planned dose and pre-RT SUV of the parotid glands, a phenomenological model of the relative change in SUV due to RT was formulated. Two models were built, one based upon planned mean dose to the parotids and mean SUV and a second model based upon dose and SUV at the voxel level in the parotid glands. Predicted and observed relative change in SUV correlated well for both the mean (Spearman's $\rho = 0.71$, $P < 0.01$) and voxel-based (Spearman's $\rho = 0.94$, $P < 0.01$) models. Subsequently, the predicted relative change in parotid SUV, which was shown to function well as an imaging biomarker for salivary toxicity in specific aim 2a, was shown to correlate closely with the observed relative change in stimulated salivary mass at the time of first follow-up. Co-registration of pre- and post-RT PET/CT images with planned dose distributions, achieved through the use of deformable image registration tools, made possible this novel dose-SUV response model. To our knowledge, this is the first time the feasibility of such a model for the human parotid glands has been demonstrated. Specific aim 2b supported the

hypothesis that predicted parotid post-RT SUV would correlate with whole mouth stimulated salivary flow ($P < 0.05$).

The central hypothesis of this work was that more accurate image registration would lead to improved quantitation of tissue response to RT with ^{18}F -FDG PET/CT. In each aim deformable image registration facilitated improved quantitation for tumor as well as normal tissue. The future of radiotherapy will most certainly see increased use of ^{18}F -FDG PET/CT imaging for quantitative treatment response monitoring. The application of deformable image registration to quantitative treatment response monitoring with ^{18}F -FDG PET/CT could have a profound impact on patient management. Accurate and early identification of residual disease may allow for more timely intervention, while the ability to quantify and predict toxicity of normal OAR might permit for individualized refinement of radiation treatment plan designs.

Appendix

Addhdrinfo

The following code will add user specified content to DICOM PET image headers, copy all unaltered content, and create a new set of images containing user generated content.

```
function Addhdrinfo

%The following script will add a DICOM tag or alter a current tag in a
%group of images

%CAUTION!!  Unitl updated for user input of tag or tag value, MUST
alter
%code!

PathName = uigetdir;
d = datestr(now,30);
mkdir(PathName,['Altered Images',d]);

%get the file list in the given directory
structFilesList = dir(PathName);
structFilesList(1:2) = []; %delete the default system directory name

%Convert struct FilesList to char array
cellFilesList=struct2cell(structFilesList);
cellFileName=cellFilesList(1,:);

%Delete folder name
TempDirFlag=cellFilesList(4,:);
TempDirFlag=cell2mat(TempDirFlag);

TempDirOrder=find(TempDirFlag);

if ~isempty(TempDirOrder)
    cellFileName(TempDirOrder)=[];
end

charFileName=char(cellFileName');
FileName=cellstr(charFileName);
```

```

%Set the duration of the for loop by the number of patients in the
selected
%directory
ptcnt=size(FileName, 1);

%Read file and alter DICOM tags
for i=1:ptcnt

    %Filter to skip files without the .dcm file extension
    TempFile=FileName{i};

    %Checks the files for the .dcm extension, if not .dcm exits loop
and
    %moves to the next file in the list
    if length(TempFile)> 4
        if ~isequal(TempFile(end-3:end), '.dcm')
            continue;
        end
    else
        continue;
    end

    %Writest the dicom header to a cell array
    X =[PathName, '\', FileName{i}];
    img=dicomread(X);
    metadata = dicominfo (X, 'dictionary','DicomDict_Plan.txt');

    %DICOM tag to be added or altered
    metadata.PatientSize = 1.65;

    %Creates a new DICOM image with altered an altered DICOM tag
    dicomwrite(img,[PathName, '\', 'Altered
Images',d, '\',FileName{i}],...
        metadata, 'CreateMode', 'copy');

end

```

Scanorder

The following script will open a series of images and place them in folders based upon chronology e.g. CT01, CT02 etc.

```

%This script will rename image directories either to CT01 or CT02 and
PET01
%or PET02
PathName = uigetdir;
%Get the file list in the given directory
FolderStruct = dir(PathName);
%Delete the default system directory name
FolderStruct(1:2) = [];

```



```

%Create structure of subdirectories

for i=1:size(FolderStruct,1)
    tempsubdir = dir([PathName, '\', FolderStruct(i,1).name]);
    tempsubdir(1:2) = [];
    if size(tempsubdir,1) > 1
        disp([FolderStruct(i,1).name, ' has more than one series']);
        continue;
    end
    structFilesList = dir([PathName, '\', FolderStruct(i,1).name, '\', ...
        tempsubdir.name]);
    %delete the default system directory name
    structFilesList(1:2) = [];
    %Convert struct FilesList to char array
    cellFilesList=struct2cell(structFilesList);
    cellFileName=cellFilesList(1,:);
    %Delete folder name
    TempDirFlag=cellFilesList(4,:);
    TempDirFlag=cell2mat(TempDirFlag);
    TempDirOrder=find(TempDirFlag);
    if ~isempty(TempDirOrder)
        cellFileName(TempDirOrder) = [];
    end
    charFileName=char(cellFileName);
    FileName=cellstr(charFileName);
    img = [PathName, '\', FolderStruct(i,1).name, '\', ...
        tempsubdir.name, '\', FileName{1}];
    info = dicominfo(img, 'dictionary', 'DicomDict_Plan.txt');
    FolderStruct(i,1).creationdate = info.InstanceCreationDate;
end

if (FolderStruct(1,1).name(1:2) == 'CT' & FolderStruct(2,1).name(1:2)
== 'CT')
    if (datenum(FolderStruct(1,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(2,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(1,1).name], [PathName, '\CT01']);

movefile([PathName, '\', FolderStruct(2,1).name], [PathName, '\CT02']);
    if (datenum(FolderStruct(3,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(4,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\PET01']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\PET02']);
    else

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\PET02']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\PET01']);
    end
    elseif (datenum(FolderStruct(1,1).creationdate, 'yyyymmdd') == ...
        datenum(FolderStruct(2,1).creationdate, 'yyyymmdd'));
        disp('Pre- and post-RT scans have same date. ');
        continue;
    else

```

```

movefile([PathName, '\', FolderStruct(1,1).name], [PathName, '\CT02']);

movefile([PathName, '\', FolderStruct(2,1).name], [PathName, '\CT01']);
    if (datenum(FolderStruct(3,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(4,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\PET01']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\PET02']);
    else

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\PET02']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\PET01']);
    end
    end
elseif (FolderStruct(1,1).name(1:2) == 'PE' &
FolderStruct(2,1).name(1:2) == 'PE')
    if (datenum(FolderStruct(1,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(2,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(1,1).name], [PathName, '\PET01']);

movefile([PathName, '\', FolderStruct(2,1).name], [PathName, '\PET02']);
    if (datenum(FolderStruct(3,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(4,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\CT01']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\CT02']);
    else

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\CT02']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\CT01']);
    end
    elseif (datenum(FolderStruct(1,1).creationdate, 'yyyymmdd') == ...
        datenum(FolderStruct(2,1).creationdate, 'yyyymmdd'));
        disp('Pre- and post-RT scans have same date. ');
        continue;
    else

movefile([PathName, '\', FolderStruct(1,1).name], [PathName, '\PET02']);

movefile([PathName, '\', FolderStruct(2,1).name], [PathName, '\PET01']);
    if (datenum(FolderStruct(3,1).creationdate, 'yyyymmdd') < ...
        datenum(FolderStruct(4,1).creationdate, 'yyyymmdd'));

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\CT01']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\CT02']);
    else

movefile([PathName, '\', FolderStruct(3,1).name], [PathName, '\CT02']);

movefile([PathName, '\', FolderStruct(4,1).name], [PathName, '\CT01']);

```

```

        end
    end
else
    disp('Not all PET/CT series present');
end
clear;

```

Pthdrinfo

The following script will extract user specified content from a series of DICOM image headers and place all desired data into an excel file.

```

function Pthdrinfo

%Extract info from DICOM headers
%Specifically for PET images
%See line 74 for list of tags captured

PathName = uigetdir;

%get the file list in the given directory
structFilesList = dir(PathName);
structFilesList(1:2) = []; %delete the default system directory name

%Convert struct FilesList to char array
cellFilesList=struct2cell(structFilesList);
cellFileName=cellFilesList(1,:);

%Delete folder name
TempDirFlag=cellFilesList(4,:);
TempDirFlag=cell2mat(TempDirFlag);

TempDirOrder=find(TempDirFlag);

if ~isempty(TempDirOrder)
    cellFileName(TempDirOrder)=[ ];
end

charFileName=char(cellFileName');
FileName=cellstr(charFileName);

%Set the duration of the for loop by the number of patients in the
selected
%directory
ptcnt=size(FileName, 1);

```

```

%Create data matrix
Ptdata=[];

%Read file and extract DICOM tags
for i=1:ptcnt

    %Filter to skip files without the .dcm file extension
    TempFile=FileName{i};

    %Checks the files for the .dcm extension, if not .dcm exits loop
and
    %moves to the next file in the list
    if length(TempFile)> 4
        if ~isequal(TempFile(end-3:end), '.dcm')
            continue;
        end
    else
        continue;
    end

    %Writest the dicom header to a cell array
    img =[PathName, '\', FileName{i}];
    info = dicominfo (img, 'dictionary','DicomDict_Plan.txt');

    %Determines if the header is from a PET image, ie
Radiopharmecuticals,
    %if not, sets those fields as zero
    if isfield(info, 'RadiopharmaceuticalInformationSequence') &&
ischar('RadiopharmaceuticalInformationSequence.Item_1.Radiopharmaceutic
alStartTime')&&
ischar('RadiopharmaceuticalInformationSequence.Item_1.RadionuclideTotal
Dose');
        Ptx =
[str2num(info.PatientID),str2num(info.StudyDate),info.PatientSize,
info.PatientWeight,info.SliceThickness,...

str2num(info.SoftwareVersion),info.ReconstructionDiameter,info.PixelSpa
cing(1),info.RescaleSlope,...
        str2num(info.AcquisitionTime),str2num(info.SeriesTime),...

str2num(info.RadiopharmaceuticalInformationSequence.Item_1.Radiopharmac
euticalStartTime),...

info.RadiopharmaceuticalInformationSequence.Item_1.RadionuclideTotalDos
e,{info.ManufacturerModelName},{info.CorrectedImage}];
    else
        Ptx =
[str2num(info.PatientID),str2num(info.StudyDate),info.PatientSize,
info.PatientWeight,info.SliceThickness,...

str2num(info.SoftwareVersion),info.ReconstructionDiameter,info.PixelSpa
cing(1),info.RescaleSlope,...

str2num(info.AcquisitionTime),str2num(info.SeriesTime),0,0,0,0];
    end
end

```

```

        %Writes tags to a matrix
        Ptdata=[Ptdata; Ptx];
    end

    %Character array that is written to excel file as label for data
    hdr = {'MRN','Date', 'Patient Height (m)', 'Patient Weight (kg)',
        'Slice Thickness (cm)', 'Software Version', 'Recon Diameter (mm)', 'Pixel
        Spacing (mm)',...
        'Rescale Slope', 'Acquisition Time (HHMMSS)', 'Series Time
        (HHMMSS)', 'RadioPharm Start Time (HHMMSS)', 'Dose (Bq)', 'Model
        Name', 'Applied Corrections'};

    %Write patient data matrix to excel file
    d = datestr(now,30);
    xlswrite([PathName, '\', 'Header Data', '_', d, '.xls'],hdr, 'Header Data');
    xlswrite([PathName, '\', 'Header Data', '_', d, '.xls'],Ptdata, 'Header
    Data', 'A2');

```

PETThreshold

The following script is designed for use with the Pinnacle³ treatment planning system to create threshold contours on PET images.

```

// Recalculate stats for max pixel
RoiList.# "Max Pixel".RecomputeStatistics="";

// Autocontour 50% threshold
Store.FloatAt.Fifty = RoiList.# "Max Pixel".MaxDisplay;
Store.At.Fifty.Multiply=0.5;
RoiList.# "50Primary".AutoLower = Store.At.Fifty.Value;
RoiList.# "50Primary".MakeCurrent="";
RoiList .# "50Primary".AutoContourROI = "Current Region Of Interest";
Store.FreeAt.Fifty="";

// Change 50% threshold data set to CT
Store.StringAt.CTdata = RoiList.# "C2".VolumeName;
RoiList.# "50Primary".VolumeName = CTdata;

// Cleanup ROI
RoiList .Current .Clean = "Rescan";
RoiList .Current .CleanAndDelete = "Delete Curves";

//Eliminate all autocontours in head
RoiExpandControl.SourceRoiList.Current.Remove="";
RoiExpandControl.AvoidRoiList.Current.Remove="";
RoiExpandControl.TargetRoi.Remove="";
RoiList.Current="50Primary";
RoiExpandControl .AddSourceRoi = "Add -->";

```

```

RoiList .Current = "Head";
RoiExpandControl .AddAvoidRoi = "Add -->";
RoiExpandControl .Expand = 0;
RoiExpandControl .TargetRoi = "na";
RoiExpandControl .DoExpand = "Proceed";

// Remove curves from 50Primaryprimary
RoiList.Current="50Primary";
RoiList.Current.RemoveAllCurves = "";
Test.ExpectAskYesNo = 1;

//Move 50% threshold curves from NA back to 50Primary
RoiExpandControl.SourceRoiList.Current.Remove="";
RoiExpandControl.AvoidRoiList.Current.Remove="";
RoiExpandControl.TargetRoi.Remove="";
RoiList .Current = "na";
RoiExpandControl .AddSourceRoi = "Add -->";
RoiExpandControl .Expand = 1;
RoiExpandControl .TargetRoi = "50Primary";
RoiExpandControl .DoExpand = "Proceed";

// Remove curves from NA
RoiList .Current = "na";
RoiList.Current.RemoveAllCurves = "";
Test.ExpectAskYesNo = 1;

```

petctfusion

The following script is a command line fusion image viewer designed for use with CT images and compressed PET images in the *.txtimg format. Of particular importance is the DualMap.m script for combined colorbars

```

function [RTstruct RTalphamap AxialImage, CTmax, CTrange,path,file] =
petctfusion(x)

%The following script will create a dual color map PET/CT figure. The
CT
%image set should be a reference and the PET overlay should be an ROI.
The
%input should be the slice number.

%Select the reference image set
[Xpixels, Ypixels, ZsliceNum, AxialImage] = OpenPinnacleImages;
%Select the overlay structure
[RTstruct,path,file] = OverlayImg(Xpixels,Ypixels,ZsliceNum);
%Perform a hard W/L on the reference CT images
[AxialImage,CTmax,CTrange] = SoftTissue(AxialImage);
%Create the alpha map to determine opacity of the overlay image

```

```

RTalphamap=RTalpha(RTstruct);
%Create the figure with a dual color map
DualMap(AxialImage, RTstruct,x,RTalphamap,CTmax,CTrange);

```

OpenPinnacleImages

The following script was utilized to open binary format images compatible with the Pinnacle³ treatment planning system.

```

function [Xpixels, Ypixels, ZsliceNum, AxialImage] =
OpenPinnacleImages(hdr)

if isdir(['G:\Research\Head and Neck\H&N PET_CT deformable
registration\...\
'Data and Analysis\Data\']);

    [file path FilterIndex]=uigetfile('*.header','Select image header
file.',...
    ['G:\Research\Head and Neck\H&N PET_CT deformable
registration\...\
'Data and Analysis\Data\']);
    hdr = [path file];
    imfile = [file(1:end-6) 'img'];
else
    [file path]=uigetfile('*.header','Select image header file. ');
    hdr = [path file];
    imfile = [file(1:end-6) 'img'];
end

%The following function will open Pinnacle CT images
HeaderInfo=textread(hdr, '%s', 'delimiter', '\n');
a=char(39);
HeaderInfo = regexp(HeaderInfo, '"', a);

%Get Byte order, actual operation
for kk=1:20
    eval(HeaderInfo{kk});
end

Xpixels=x_dim;
Ypixels=y_dim;
ZsliceNum=z_dim;

%Read daily pinnacle data
if byte_order == 0
    fid=fopen([path imfile], 'r', 'ieee-le');
else

```

```

        fid=fopen([path imfile], 'r', 'ieee-be');
    end

    [TempData, Count]=fread(fid, Xpixels*Ypixels*ZsliceNum, '*int16');
    TempData=uint16(TempData);
    fclose(fid);

    AxialImage=reshape(TempData, [Xpixels, Ypixels, ZsliceNum]);
    AxialImage=permute(AxialImage, [2,1, 3]);

```

OverlayImg

The following code will turn the compressed image files in the *.txtimg format into false color images that can be fused with grayscale CT images.

```

function [RTstruct,path,file] = OverlayImg(Xpixels,Ypixels,ZsliceNum)

%This code will turn the *.txtimg files into useable images based upon
%outputs from the OpenPinnacleImages function

if isdir(['C:\Documents and Settings\bacannon\My
Documents\Research\']);

    [file path FilterIndex]=uigetfile('*.txtimg','Select text image
file.',...
    ['C:\Documents and Settings\bacannon\My Documents\Research\']);
    img = [path file];
    %imfile = [file(1:end-6) 'img'];
else
    [file path]=uigetfile('*.txtimg','Select text image file. ');
    img = [path file];
    %imfile = [file(1:end-6) 'img'];
end

fid = fopen(img,'r');
%The following function will open the text image file
Temp = textscan(fid,'% -15.0f% -15.0f% -15.0f% -15.4f% -15.4f% -15.4f% -
15.4f');
ImgData = [Temp{1},Temp{2},Temp{3},Temp{4},Temp{5},Temp{6},Temp{7}];
clear Temp;
fclose(fid);

%Create the empty matrix that will be the image
RTstruct = zeros (Xpixels, Ypixels, ZsliceNum);
%Determine the indices for the data in the volumetric data set
IND = sub2ind(size(RTstruct),ImgData(:,2),ImgData(:,1),ImgData(:,3));

%Place value of interest as pixel value

```



```

    %Column 1 - 3: x, y, and z location
    %Column 4: Fractional uptake
    %Column 5: Mean metabolic dose
    %Column 6: Mean SUV1
    %Column 7: Mean SUV2
    %By changing the data one must alter line 55 of DualMap.m
    RTstruct(IND) = ImgData(:,6);

clear IND ImgData

```

SoftTissue

The following script will perform a hard window/level operation to set the dynamic range of CT images to a soft tissue window.

```

function [AxialImage,CTmax,CTrange] = SoftTissue(AxialImage)

%This function will rescale the CT data to be in the soft tissue window
%range
CTmax = int16(1200);
CTrange = int16(400);
X = size(AxialImage,1);
Y = size(AxialImage,2);
Z = size(AxialImage,3);

%Rescal the CT image to the soft tissue window
for i=1:(X*Y*Z)
    if AxialImage(i) < 800;
        AxialImage(i)=800;
    elseif AxialImage(i) > 1200;
        AxialImage(i)=1200;
    else
        end
end
end

```

RTalpha

The code given below was utilized to create an opacity mask for the overlay images.

```

function RTalphamap=RTalpha(RTstruct)

```

```

X = size(RTstruct,1);
Y = size(RTstruct,2);
Z = size(RTstruct,3);
RTalphamap = zeros (X,Y,Z);

for i=1:(X*Y*Z)
    if RTstruct(i)==0
        RTalphamap(i)= 0;
    else
        RTalphamap(i)=1;
    end
end

```

DualMap

The following script is the most critical for image overlay in the MATLAB workspace.

```

function DualMap(AxialImage, RTstruct,x,RTalphamap,CTmax,CTrange)

%The following script will create a dual color map image with fused
%AxialImage and overlay RTstruct

%Create the figure
hf = figure('units','normalized','position',[.2 .2 .6
.6],'Color','none');
%Assign bg image to im1
im1=AxialImage(:,:,x);
%Assign fg image to im2
im2=RTstruct(:,:,x);

%Open the image
hax1 = imagesc(im1);
%Set the aspect ratio
axis('image')
axis off
hold on

%Set the colormap for the bg as grayscale
cmap1 = colormap(gray);
ax1 = gca;
set(ax1,'Color','none')
%Set the fg axes
ax2 =axes('Position',get(ax1,'Position'));
%Open the fg images and get its handle
hax2 = imagesc(im2);
%Set the aspect ratio
axis('image')

```

```

axis off
%Overlay the ax2 ontop of ax1
set(ax2,'YAxisLocation','right','Color','none','XTickLabel',[]);
set(ax2,'XLim',get(ax1,'XLim'),'Layer','top')

%Set the colormap for the fg image as jet
cmap2=colormap(jet);
%Splice the two colormaps into one
colormap([cmap1;cmap2]);
%Determine the length of the entire colormap
CmLength = length(colormap);
%Set the beginning position of the colormap
BeginSlot1 = 1;
%Set the end position of the first colormap scale
EndSlot1 = length(cmap1);
%Set the beginning position of the second colormap scale
BeginSlot2 = EndSlot1 + 1;
%Set the end position of the second colormap scale
EndSlot2 = CmLength;
%Determine the current color limits for the bg image
CLim1 = [min(min(min(AxialImage))) max(max(max(AxialImage)))];
%Determine the current color limits for the fg image
CLim2 = [min(min(min(RTstruct))) max(max(max(RTstruct)))];

%Set the CLim2 for presentation plots, same scale
CLim2 = [.4 2.0];

%Set the color limits for the bg image
set(ax1,'CLim',newclim(BeginSlot1,EndSlot1,CLim1(1),...
    CLim1(2),CmLength));
%Set the color limits for the fg image
set(ax2,'CLim',newclim(BeginSlot2,EndSlot2,CLim2(1),...
    CLim2(2),CmLength));
%Set the transparency
set(hax2, 'AlphaData',RTalphamap(:,:,x));
%Set the position of the color
cbar_axes = colorbar('Position',[.835 .11 .062 .815]);
%Set the CT label for the colorbar
Ylabel = str2num(get(cbar_axes,'YTickLabel'));
f=int16(1);
if rem(size(Ylabel,1),2)==1
    %Number of labels
    LabCnt = int16(round(size(Ylabel,1)/2)-1);
    %The colorbar has an odd number of labels
    for i=LabCnt:-1:1
        Ylabel(i,1)= CTmax - (CTrange/LabCnt)*f;
        f=f+1;
    end
else
    %Number of labels
    LabCnt = int16(size(Ylabel,1)/2);
    %The colorbar has an even number of labels
    for i=LabCnt:-1:1
        Ylabel(i,1) = CTmax - (CTrange/(LabCnt+1))*f;
        f=f+1;
    end
end

```

```
end
set(cbar_axes, 'YTickLabel', num2str(Ylabel), 'YColor', [.968 .968 .968]);
```

Metabolic Dose

The following script was utilized to construct voxel based dose-metabolic response curve.

```
function MetabolicDose_Batch(PtData, PtName, PathName, drun, Norm)

%%!!!!!! MUST!!! be used with Main_Head_Batch.m or define the global
%%                                     variables below!!!!!!

%The following algorithm will create a dose-uptake response curve for
the
%structures contained within the *.roi file based upon the metabolic
dose
%metric. Bins are created by binning the dose to achieve uniformity
and
%the computing the metabolic dose metric and fractional uptake for said
%regions.

%Define the type of analysis
%BinType = 1 - Dose binned analysis
%BinType = 2 - Function binned analysis
%BinType = 3 - Metabolic dose binned analysis
BinType = 3;
SumFlag = 0;
if isdir([PathName, '\', 'MetabolicDoseBin_data'])
    d = datestr(now, 30);
    mkdir(PathName, [ 'MetabolicDoseBin_data_', d]);
    mkdir(PathName, [ 'MetabolicDoseBin_plot_', d]);
    mkdir(PathName, [ 'MetabolicDoseBin_txtimg_', d]);
    dataWritePath = [PathName, '\', 'MetabolicDoseBin_data_', d];
    plotWritePath = [PathName, '\', 'MetabolicDoseBin_plot_', d];
    txtimgWritePath = [PathName, '\', 'MetabolicDoseBin_txtimg_', d];
else
    mkdir(PathName, 'MetabolicDoseBin_data');
    mkdir(PathName, 'MetabolicDoseBin_plot');
    mkdir(PathName, 'MetabolicDoseBin_txtimg');
    dataWritePath = [PathName, '\', 'MetabolicDoseBin_data'];
    plotWritePath = [PathName, '\', 'MetabolicDoseBin_plot'];
    txtimgWritePath = [PathName, '\', 'MetabolicDoseBin_txtimg'];
end

%Define the header string that all files will use in this function
```

```

HeaderStr={'Patient:  ' ; ...
'Dose_min(cGy) ' ; ...
'Dose_max(cGy) ' ; ...
'Dose_mean(cGy) ' ; ...
'Dose_std(cGy) ' ; ...
'SUV1_min' ; ...
'SUV1_max' ; ...
'SUV1_mean' ; ...
'SUV1_std' ; ...
'SUV2_min' ; ...
'SUV2_max' ; ...
'SUV2_mean' ; ...
'SUV2_std' ; ...
'Metabolic_Dose_min(cGyMBq/mL) ' ; ...
'Metabolic_Dose_max(cGyMBq/mL) ' ; ...
'Metabolic_Dose_mean(cGyMBq/mL) ' ; ...
'Metabolic_Dose_median(cGyMBq/mL) ' ; ...
'Metabolic_Dose_std(cGyMBq/mL) ' ; ...
'Lilliefors_MetabolicDose' ; ...
'Kurtosis_MetabolicDose' ; ...
'Skewness_MetabolicDose' ; ...
'Fractional_Uptake_min' ; ...
'Fractional_Uptake_max' ; ...
'Fractional_Uptake_mean' ; ...
'Fractional_Uptake_std' ; ...
'Lilliefors_FractionalUptake' ; ...
'Kurtosis_FractionalUptake' ; ...
'Skewness_FractionalUptake' ; ...
'VolumeFraction' ; ...
'VoxelCount' };

%This code will cause problems if not used with Main_Head_Batch
TempIndex=strmatch('Patient:  ', HeaderStr);
HeaderStr{TempIndex}=[ 'Patient:  ', PtName];

%Gather all ROIs present in PtData structure
FieldNames=fieldnames(PtData);

for j=1:length(FieldNames)
    %disp([PtName, ' ',FieldNames{j}])
    tempFieldName = FieldNames{j};
    %Sort the ROI by physical dose
    PtData.(FieldNames{j}) = sortrows(PtData.(FieldNames{j}),4);
    %Identify voxels that received less than 500 cGy
    DoseFilter = find(PtData.(FieldNames{j})(:,4)<500);
    %Remove voxels receiving less than 500 cGy from analysis
    PtData.(FieldNames{j})(DoseFilter(:,1),:) = [];
    %If all physical dose is below threshold, do not analyze
    if isempty(PtData.(FieldNames{j}));
        continue;
    end
    %Compute the Metabolic dose metric for each voxel
    PtData.(FieldNames{j})(:,8) =
    PtData.(FieldNames{j})(:,4).*PtData.(FieldNames{j})(:,5);
end

```

```

        if
        (strcmp(tempFieldName, 'ROIGTVprimary') || strcmp(tempFieldName, 'ROI40Primary') || ...

        strcmp(tempFieldName, 'ROI50Primary') || strcmp(tempFieldName, 'ROI40PAirlmtd') || ...

        strcmp(tempFieldName, 'ROI50PAirlmtd') || strcmp(tempFieldName, 'ROI40Node') || ...

        strcmp(tempFieldName, 'ROI50Node') || strcmp(tempFieldName, 'ROI40NAirlmtd') || ...

        strcmp(tempFieldName, 'ROI50NAirlmtd') || strcmp(tempFieldName, 'ROIGTVAirilmtd') || ...

        strcmp(tempFieldName, 'ROIGTVNodesAir') || strcmp(tempFieldName, 'ROIGTVnode') || ...
            strcmp(tempFieldName, 'ROIRefBrainstem'));

        continue;
        if ((max (PtData.(FieldNames{j}(:,8)) - min
        (PtData.(FieldNames{j}(:,8)))>100000);
            DoseBin = (min (PtData.(FieldNames{j}(:,8)):10000:max
        (PtData.(FieldNames{j}(:,8)))));
        else
            DoseBin = (min (PtData.(FieldNames{j}(:,8)):2000:max
        (PtData.(FieldNames{j}(:,8)))));
        end
        elseif
        (strcmp(tempFieldName, 'ROIParotidR') || strcmp(tempFieldName, 'ROIParotidL') || ...

        strcmp(tempFieldName, 'ROIipsilateralParotid') || strcmp(tempFieldName, 'ROIcontralateralParotid') || ...

        strcmp(tempFieldName, 'ROIParotids') || strcmp(tempFieldName, 'ROISMG'))
            DoseBin = (min (PtData.(FieldNames{j}(:,8)):2000:max
        (PtData.(FieldNames{j}(:,8)))));
        else
            continue;
            if strcmp(tempFieldName, 'ROI70Gy')
                if ((max (PtData.(FieldNames{j}(:,8)) - min
        (PtData.(FieldNames{j}(:,8)))>100000);
                    DoseBin = (min
        (PtData.(FieldNames{j}(:,8)):10000:max (PtData.(FieldNames{j}(:,8)))));
                else
                    DoseBin = (min
        (PtData.(FieldNames{j}(:,8)):2000:max (PtData.(FieldNames{j}(:,8)))));
                end
            else
                continue;
            end
        %         if ((max (PtData.(FieldNames{j}(:,8)) - min
        (PtData.(FieldNames{j}(:,8)))<500)
        %             %Dose bin size set at 10 cGy

```

```

%           DoseBin = (min (PtData.(FieldNames{j})(:,8)):10:max
(PtData.(FieldNames{j})(:,8)));
%           %elseif ((max (PtData.(FieldNames{j})(:,8)) - min
(PtData.(FieldNames{j})(:,8)))>)
%           else
%           %Dose bin size set at 500 cGy
%           DoseBin = (min (PtData.(FieldNames{j})(:,8)):500:max
(PtData.(FieldNames{j})(:,8)));
%           end
end

%Create the dose bin label i.e., make the bins integers and
eliminate
%trailing and leading zeros
DoseName = strtrim(cellstr(num2str(uint32(DoseBin'))));
%Compute the total number of pixels in the ROI for fractional
volume
%purposes
VolumeTotal = size(PtData.(FieldNames{j}),1);

%Open the text file for writing fDVH data
fid = fopen([dataWritePath,'\ ', FieldNames{j},'.out'],'wt');

if
(strcmp(tempFieldName,'ROIGTVprimary')||strcmp(tempFieldName,'ROI40Prim
ary'))||...

strcmp(tempFieldName,'ROI50Primary')||strcmp(tempFieldName,'ROI40PAirlm
td'))||...

strcmp(tempFieldName,'ROI50PAirlmtd')||strcmp(tempFieldName,'ROIGTVairl
mtd'))||...

strcmp(tempFieldName,'ROIContralateralParotid')||strcmp(tempFieldName,'
ROI IpsilateralParotid'))||...

strcmp(tempFieldName,'ROIParotidR')||strcmp(tempFieldName,'ROIParotidL'
))||...
    strcmp(tempFieldName,'ROIParotids'));

    %Open file for writing text image data
    fid2 = fopen([txtimgWritePath,'\ ',
tempFieldName,'.txtimg'],'wt');
end

%Write the header to text file
for k=1:length(HeaderStr)
    if k == 1
        fprintf(fid, '%-35s\n\n', HeaderStr{k});
    elseif k == (length(HeaderStr))
        fprintf(fid, '%-35s\n', HeaderStr{k});
    else
        fprintf(fid, '%-35s', HeaderStr{k});
    end
end
%Initialize the text image matrix

```

```

    ImgTxtData = [];
    BinData = zeros(29,size(DoseBin,2));
    for i=1:(size(DoseBin,2));
        MetDose = [];
        %Identify indecies of data falling into volume-dose bins
        if i == size(DoseBin,2);
            %Finds indicies for the last dose bin
            PtHisto.(FieldNames{j}).DoseBin{i} =
find((PtData.(FieldNames{j}))(:,8) >= DoseBin(i));
        else
            %Finds indicies for all bins except last bin
            PtHisto.(FieldNames{j}).DoseBin{i} =
find((PtData.(FieldNames{j}))(:,8) >= DoseBin(i) &
(PtData.(FieldNames{j}))(:,8) < DoseBin(i+1)));
        end

        %Extract the needed data to individual volume-dose bins
        PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]) =
PtData.(FieldNames{j})(PtHisto.(FieldNames{j}).DoseBin{1,i},:);
        %If the dose bin is empty, simply set datum as NaN
        if isempty(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,7));
            BinData(1:29,i) = NaN;
            continue;
        elseif size(PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]),1)
< 4;
            %Write the Dose data to matrix
            BinData(1,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,4));
            BinData(2,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,4));
            BinData(3,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,4));
            BinData(4,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,4));
            %Write SUV1 data to matrix
            BinData(5,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,5));
            BinData(6,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,5));
            BinData(7,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,5));
            BinData(8,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,5));
            %Write SUV2 data to matrix
            BinData(9,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,6));
            BinData(10,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,6));
            BinData(11,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,6));
            BinData(12,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,6));
            %Write Metabolic Dose data to matrix
            MetDose = PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}])(:,4) .* ...;
            PtHisto.(FieldNames{j}).data.(['V' DoseName{i}])(:,5);

```



```

        BinData(13,i) = min(MetDose);
        BinData(14,i) = max(MetDose);
        BinData(15,i) = mean(MetDose);
        BinData(16,i) = median(MetDose);
        BinData(17,i) =
((mean(MetDose)^2)*((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))/...;
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5)))^2)+...;
        (var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4))/...;
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4)))^2)))^0.5;
        %Do not perform Lilliefors
        BinData(18,i) = NaN;
        %Do not perform Kurtosis
        BinData(19,i) = NaN;
        %Do not perform Skewness
        BinData(20,i) = NaN;
        %Write Fractional Uptake data to matrix
        BinData(21,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        BinData(22,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        BinData(23,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        %Write the std of F.U., using propagation of error with
        %correlated data term
        FU = cov(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5),...
        PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]))(:,6));
        BinData(24,i) = ((mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7)))^2*...
        ((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5)))^2)+...
        ((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6)))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6)))^2)-...
        (2*FU(1,2)/(mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))*...
        mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6))))))^0.5;
        %Do not perform Lilliefors
        BinData(25,i) = NaN;
        %Do not perform Kurtosis
        BinData(26,i) = NaN;
        %Do not perform Skewness
        BinData(27,i) = NaN;
        %Write fractional volume of current bin to matrix
        BinData(28,i) = size(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]),1)/VolumeTotal;
        %Write number of voxels of current bin to matrix
        BinData(29,i) = size(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]),1);

```

```

        %Do not perform tests for normality of bin

    else
        %Write the Dose data to matrix
        BinData(1,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4));
        BinData(2,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4));
        BinData(3,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4));
        BinData(4,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4));
        %Write SUV1 data to matrix
        BinData(5,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5));
        BinData(6,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5));
        BinData(7,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5));
        BinData(8,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5));
        %Write SUV2 data to matrix
        BinData(9,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6));
        BinData(10,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6));
        BinData(11,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6));
        BinData(12,i) = std(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6));
        %Write Metabolic Dose data to matrix
        MetDose = PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4) .* ...;
        PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]))(:,5);
        BinData(13,i) = min(MetDose);
        BinData(14,i) = max(MetDose);
        BinData(15,i) = mean(MetDose);
        BinData(16,i) = median(MetDose);
        BinData(17,i) =
        (mean(MetDose)^2*((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))^2)+...
        (var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,4))^2))^0.5;
        %Perform Lilliefors test for goodness of fit to normal
        %distribution to metabolic dose data
        BinData(18,i) = lillietest(MetDose);
        %Determine sample kurtosis, normal = 3, more outlier
prone>3,
        %less outlier prone<3 for metabolic dose data
        BinData(19,i) = kurtosis(MetDose);
        %Determine sample skewness, if negative data are spread to
left
        %of mean, if positive data are to right of mean, zero is

```

```

        %symmetric distribution
        BinData(20,i) = skewness(MetDose);
        %Write Fractional Uptake data to matrix
        BinData(21,i) = min(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        BinData(22,i) = max(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        BinData(23,i) = mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        %Write the std of F.U., using propagation of error with
        %correlated data term
        FU = cov(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5),...
        PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]))(:,6));
        BinData(24,i) = ((mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7)))^2*...
        ((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5)))^2)+...
        ((var(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6)))/...
        (mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6)))^2)-...
        (2*FU(1,2)/(mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,5))*...
        mean(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,6))))^0.5;
        %Perform Lilliefors test for goodness of fit to normal
        %distribution
        BinData(25,i) =
lillietest(PtHisto.(FieldNames{j}).data.(['V' DoseName{i}]))(:,7));
        %Determine sample kurtosis, normal = 3, more outlier
prone>3,
        %less outlier prone<3
        BinData(26,i) = kurtosis(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        %Determine sample skewness, if negative data are spread to
left
        %of mean, if positive data are to right of mean, zero is
        %symmetric distribution
        BinData(27,i) = skewness(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,7));
        %Write fractional volume of current bin to matrix
        BinData(28,i) = size(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]),1)/VolumeTotal;
        %Write number of voxels of current bin to matrix
        BinData(29,i) = size(PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]),1);
    end
    if
    (strcmp(tempFieldName,'ROIGTVprimary')||strcmp(tempFieldName,'ROI40Prim
ary'))||...
    strcmp(tempFieldName,'ROI50Primary')||strcmp(tempFieldName,'ROI40PAirlm
td'))||...

```

```

strcmp(tempFieldName, 'ROI50PAirlmtd') || strcmp(tempFieldName, 'ROI50PAirlmtd') || ...

strcmp(tempFieldName, 'ROIContralateralParotid') || strcmp(tempFieldName, 'ROIContralateralParotid') || ...

strcmp(tempFieldName, 'ROIParotidR') || strcmp(tempFieldName, 'ROIParotidL') || ...

        strcmp(tempFieldName, 'ROIParotids'));

        tempMetDoseName =
char(strtrim(cellstr(num2str(uint32(BinData(15,i))))));
        tempFracUptkName =
char(strtrim(cellstr(num2str(uint32(BinData(23,i))))));
        tempdosename = char(DoseName(i));
        %Create histogram of dose bin
        SubPlotohisto(tempFieldName, tempMetDoseName,
tempFracUptkName, tempdosename, PtHisto, plotWritePath, MetDose, Norm,
SumFlag);

        clear tempdosename tempMetDoseName tempFracUptkName FU
        %Create temporary matrix of current dose bin's x, y, and z
        indices
        tempTxtData = (PtHisto.(FieldNames{j}).data.(['V'
DoseName{i}]))(:,1:3));
        %Place the dose bin's mean fractional uptake as the pixel
        value
        tempTxtData(4,:) = BinData(23,i);
        %Place the mean metabolic dose for this region
        tempTxtData(5,:) = BinData(15,i);
        %Place the mean SUV1 for this region
        tempTxtData(6,:) = BinData(7,i);
        %Place the mean SUV2 for this region
        tempTxtData(7,:) = BinData(11,i);
        %Write image data to text file
        fprintf(fid2, '%-15.0f%-15.0f%-15.0f%-15.4f%-15.4f%-15.4f%-
15.4f\n', tempTxtData);
        tempTxtData=[];
    end
end
%Eliminate the NaNs from the BinData matrix
EmptyBinIndex = isnan(BinData(1,:));
BinData(:,EmptyBinIndex) = [];
%Write fDVH plot to file calling subfunction
SubPlotfDVH(tempFieldName, BinData, plotWritePath, Norm, BinType);
%Write differential dose volume histogram
SubPlotdiffDVH(tempFieldName, BinData, plotWritePath, Norm,
BinType);
%Write cummulative dose volume histogram
SubPlotcummdVH(tempFieldName, BinData, plotWritePath, Norm,
BinType);
    if
        (strcmp(tempFieldName, 'ROI50Primary') || strcmp(tempFieldName, 'ROI40Prim
ary')) || ...

        strcmp(tempFieldName, 'ROI50Primary') || strcmp(tempFieldName, 'ROI40PAirlm
td') || ...

```

```

strcmp(tempFieldName, 'ROI50PAirlmtd') || strcmp(tempFieldName, 'ROIGTVAir1
mtd') || ...

strcmp(tempFieldName, 'ROIContralateralParotid') || strcmp(tempFieldName, '
ROIipsilateralParotid') || ...

strcmp(tempFieldName, 'ROIParotidR') || strcmp(tempFieldName, 'ROIParotidL'
) || ...
        strcmp(tempFieldName, 'ROIParotids'));

        %Close the text image
        fclose(fid2);
    end
    %Write the BinData matrix to
    fprintf(fid, '%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-
35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-
35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-35.4f%-
35.4f%-35.4f\n', BinData);
    fclose(fid);
    %Write data to summary file for entire patient cohort
    summaryfile(tempFieldName, BinData, PathName, drun, PtName,
BinType, PtData);
    %Write unbinmed data to summary file for entire cohort
    %Unbinnedsummaryfile(tempFieldName, PtData, PathName, drun, PtName,
BinType);
    clear PtHisto DoseBin DoseName VolumeTotal BinData tempFieldName
tempTxtData;
end

```

References

1. Giraud, P., Grahek, D., Montravers, F., Carette, M.F., Deniaud-Alexandre, E., Julia, F., Rosenwald, J.C., Cosset, J.M., Talbot, J.N., Housset, M., and Touboul, E., *CT and (18)F-deoxyglucose (FDG) image fusion for optimization of conformal radiotherapy of lung cancers*. International journal of radiation oncology, biology, physics, 2001. **49**(5): p. 1249-57.
2. Hopper, K.D., Singapuri, K., and Finkel, A., *Body CT and oncologic imaging*. Radiology, 2000. **215**(1): p. 27-40.
3. Jones, D.R., Parker, L.A., Jr., Detterbeck, F.C., and Egan, T.M., *Inadequacy of computed tomography in assessing patients with esophageal carcinoma after induction chemoradiotherapy*. Cancer, 1999. **85**(5): p. 1026-32.
4. Tenhunen, M., Collan, J., Kouri, M., Kangasmaki, A., Heikkonen, J., Kairemo, K., Makitie, A., Joensuu, H., and Saarilahti, K., *Scintigraphy in prediction of the salivary gland function after gland-sparing intensity modulated radiation therapy for head and neck cancer*. Radiother Oncol, 2008. **87**(2): p. 260-7.
5. Therasse, P., Arbuck, S.G., Eisenhauer, E.A., Wanders, J., Kaplan, R.S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A.T., Christian, M.C., and Gwyther, S.G., *New guidelines to evaluate the response to treatment in solid tumors. European Organization for*

Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst, 2000. 92(3): p. 205-16.

6. Laubenbacher, C., Saumweber, D., Wagner-Manslau, C., Kau, R.J., Herz, M., Avril, N., Ziegler, S., Kruschke, C., Arnold, W., and Schwaiger, M., *Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous-cell carcinomas. J Nucl Med, 1995. 36(10): p. 1747-57.*
7. Ojiri, H., Mancuso, A.A., Mendenhall, W.M., and Stringer, S.P., *Lymph nodes of patients with regional metastases from head and neck squamous cell carcinoma as a predictor of pathologic outcome: size changes at CT before and after radiation therapy. AJNR Am J Neuroradiol, 2002. 23(10): p. 1627-31.*
8. Gayed, I., Vu, T., Iyer, R., Johnson, M., Macapinlac, H., Swanston, N., and Podoloff, D., *The role of 18F-FDG PET in staging and early prediction of response to therapy of recurrent gastrointestinal stromal tumors. J Nucl Med, 2004. 45(1): p. 17-21.*
9. Bar-Shalom, R., Yefremov, N., Guralnik, L., Gaitini, D., Frenkel, A., Kuten, A., Altman, H., Keidar, Z., and Israel, O., *Clinical performance of PET/CT in evaluation of cancer: additional value for diagnostic imaging and*

- patient management*. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2003. **44**(8): p. 1200-9.
10. Daisne, J.F., Sibomana, M., Bol, A., Doumont, T., Lonneux, M., and Grêgoire, V., *Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms*. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology, 2003. **69**(3): p. 247-50.
 11. Paulino, A.C., Koshy, M., Howell, R., Schuster, D., and Davis, L.W., *Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer*. Int J Radiat Oncol Biol Phys, 2005. **61**(5): p. 1385-92.
 12. Hautzel, H. and Mèuller-Gèartner, H.W., *Early changes in fluorine-18-FDG uptake during radiotherapy*. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 1997. **38**(9): p. 1384-6.
 13. Humm, J.L., Lee, J., O'Donoghue, J.A., Squire, O., Ling, C.C., Pentlow, K., Mehta, B., Erdi, Y., Ruan, S., and Larson, S.M., *Changes in FDG Tumor Uptake during and after Fractionated Radiation Therapy in a Rodent Tumor Xenograft*. Clin Positron Imaging, 1999. **2**(5): p. 289-296.
 14. Schwartz, D.L., Ford, E., Rajendran, J., Yueh, B., Coltrera, M.D., Virgin, J., Anzai, Y., Haynor, D., Lewellyn, B., Mattes, D., Meyer, J., Phillips, M.,

- Leblanc, M., Kinahan, P., Krohn, K., Eary, J., and Laramore, G.E., *FDG-PET/CT imaging for preradiotherapy staging of head-and-neck squamous cell carcinoma*. International journal of radiation oncology, biology, physics, 2005. **61**(1): p. 129-36.
15. Schwartz, D.L., Ford, E.C., Rajendran, J., Yueh, B., Coltrera, M.D., Virgin, J., Anzai, Y., Haynor, D., Lewellen, B., Mattes, D., Kinahan, P., Meyer, J., Phillips, M., Leblanc, M., Krohn, K., Eary, J., and Laramore, G.E., *FDG-PET/CT-guided intensity modulated head and neck radiotherapy: a pilot investigation*. Head Neck, 2005. **27**(6): p. 478-87.
 16. Schwartz, D.L., Rajendran, J., Yueh, B., Coltrera, M., Anzai, Y., Krohn, K., and Eary, J., *Staging of head and neck squamous cell cancer with extended-field FDG-PET*. Archives of otolaryngology--head & neck surgery, 2003. **129**(11): p. 1173-8.
 17. Schwartz, D.L., Rajendran, J., Yueh, B., Coltrera, M.D., Leblanc, M., Eary, J., and Krohn, K., *FDG-PET prediction of head and neck squamous cell cancer outcomes*. Archives of otolaryngology--head & neck surgery, 2004. **130**(12): p. 1361-7.
 18. Weber, W.A., Avril, N., and Schwaiger, M., *Relevance of positron emission tomography (PET) in oncology*. Strahlentherapie und Onkologie : Organ der Deutschen Röntgengesellschaft ... [et al], 1999. **175**(8): p. 356-73.

19. Gambhir, S.S., Czernin, J., Schwimmer, J., Silverman, D.H., Coleman, R.E., and Phelps, M.E., *A tabulated summary of the FDG PET literature*. J Nucl Med, 2001. **42**(5 Suppl): p. 1S-93S.
20. Veit-Haibach, P., Luczak, C., Wanke, I., Fischer, M., Egelhof, T., Beyer, T., Dahmen, G., Bockisch, A., Rosenbaum, S., and Antoch, G., *TNM staging with FDG-PET/CT in patients with primary head and neck cancer*. Eur J Nucl Med Mol Imaging, 2007. **34**(12): p. 1953-62.
21. Kasamon, Y.L. and Wahl, R.L., *FDG PET and risk-adapted therapy in Hodgkin's and non-Hodgkin's lymphoma*. Curr Opin Oncol, 2008. **20**(2): p. 206-19.
22. Higashi, K., Clavo, A.C., and Wahl, R.L., *In vitro assessment of 2-fluoro-2-deoxy-D-glucose, L-methionine and thymidine as agents to monitor the early response of a human adenocarcinoma cell line to radiotherapy*. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 1993. **34**(5): p. 773-9.
23. Patlak, C.S., Blasberg, R.G., and Fenstermacher, J.D., *Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data*. J Cereb Blood Flow Metab, 1983. **3**(1): p. 1-7.
24. Huang, S.C., *Anatomy of SUV. Standardized uptake value*. Nuclear medicine and biology, 2000. **27**(7): p. 643-6.

25. Graham, M.M., Peterson, L.M., and Hayward, R.M., *Comparison of simplified quantitative analyses of FDG uptake*. Nucl Med Biol, 2000. **27**(7): p. 647-55.
26. Greven, K.M., Williams, D.W., 3rd, Keyes, J.W., Jr., McGuirt, W.F., Watson, N.E., Jr., Randall, M.E., Raben, M., Geisinger, K.R., and Cappellari, J.O., *Positron emission tomography of patients with head and neck carcinoma before and after high dose irradiation*. Cancer, 1994. **74**(4): p. 1355-9.
27. Allal, A.S., Dulguerov, P., Allaoua, M., Haenggeli, C.A., El-Ghazi el, A., Lehmann, W., and Slosman, D.O., *Standardized uptake value of 2-[(18)F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy*. J Clin Oncol, 2002. **20**(5): p. 1398-404.
28. Kitagawa, Y., Sadato, N., Azuma, H., Ogasawara, T., Yoshida, M., Ishii, Y., and Yonekura, Y., *FDG PET to evaluate combined intra-arterial chemotherapy and radiotherapy of head and neck neoplasms*. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 1999. **40**(7): p. 1132-7.
29. Kao, J., Vu, H.L., Genden, E.M., Mocherla, B., Park, E.E., Packer, S., Som, P.M., and Kostakoglu, L., *The diagnostic and prognostic utility of*

- positron emission tomography/computed tomography-based follow-up after radiotherapy for head and neck cancer. Cancer, 2009.*
30. Yao, M., Smith, R.B., Hoffman, H.T., Funk, G.F., Lu, M., Menda, Y., Graham, M.M., and Buatti, J.M., *Clinical significance of postradiotherapy [18F]-fluorodeoxyglucose positron emission tomography imaging in management of head-and-neck cancer-a long-term outcome report. Int J Radiat Oncol Biol Phys, 2009. 74(1): p. 9-14.*
 31. Arslan, N., Miller, T.R., Dehdashti, F., Battafarano, R.J., and Siegel, B.A., *Evaluation of response to neoadjuvant therapy by quantitative 2-deoxy-2-[18F]fluoro-D-glucose with positron emission tomography in patients with esophageal cancer. Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging, 2002. 4(4): p. 301-10.*
 32. Yen, T.C., Lin, C.Y., Wang, H.M., Huang, S.F., Liao, C.T., Kang, C.J., Ng, S.H., Chan, S.C., Fan, K.H., Chen, I.H., Lin, W.J., Cheng, A.J., and Chang, J.T., *18F-FDG-PET for evaluation of the response to concurrent chemoradiation therapy with intensity-modulated radiation technique for Stage T4 nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys, 2006. 65(5): p. 1307-14.*
 33. Greven, K.M., Williams, D.W., 3rd, McGuirt, W.F., Sr., Harkness, B.A., D'Agostino, R.B., Jr., Keyes, J.W., Jr., and Watson, N.E., Jr., *Serial*

- positron emission tomography scans following radiation therapy of patients with head and neck cancer.* Head Neck, 2001. **23**(11): p. 942-6.
34. Moeller, B.J., Rana, V., Cannon, B.A., Williams, M.D., Sturgis, E.M., Ginsberg, L.E., Macapinlac, H.A., Lee, J.J., Ang, K.K., Chao, K.S., Chronowski, G.M., Frank, S.J., Morrison, W.H., Rosenthal, D.I., Weber, R.S., Garden, A.S., Lippman, S.M., and Schwartz, D.L., *Prospective risk-adjusted [18F]Fluorodeoxyglucose positron emission tomography and computed tomography assessment of radiation response in head and neck cancer.* J Clin Oncol, 2009. **27**(15): p. 2509-15.
 35. Falen, S., Ivanovic, M., and Venkataraman, S., *Standard Uptake Value analysis: evaluation of acquisition and processing parameters for PET.* Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging, 2003. **5**: p. 106.
 36. Boellaard, R., Krak, N.C., Hoekstra, O.S., and Lammertsma, A.A., *Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study.* Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2004. **45**(9): p. 1519-27.
 37. Hoekstra, C.J., Paglianiti, I., Hoekstra, O.S., Smit, E.F., Postmus, P.E., Teule, G.J., and Lammertsma, A.A., *Monitoring response to therapy in cancer using [18F]-2-fluoro-2-deoxy-D-glucose and positron emission*

- tomography: an overview of different analytical methods.* European journal of nuclear medicine, 2000. **27**(6): p. 731-43.
38. Young, H., Baum, R., Cremerius, U., Herholz, K., Hoekstra, O., Lammertsma, A.A., Pruim, J., and Price, P., *Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group.* Eur J Cancer, 1999. **35**(13): p. 1773-82.
 39. Chao, K.S., Deasy, J.O., Markman, J., Haynie, J., Perez, C.A., Purdy, J.A., and Low, D.A., *A prospective study of salivary function sparing in patients with head-and-neck cancers receiving intensity-modulated or three-dimensional radiation therapy: initial results.* Int J Radiat Oncol Biol Phys, 2001. **49**(4): p. 907-16.
 40. Guchelaar, H.J., Vermes, A., and Meerwaldt, J.H., *Radiation-induced xerostomia: pathophysiology, clinical course and supportive treatment.* Support Care Cancer, 1997. **5**(4): p. 281-8.
 41. Cooper, J.S., Fu, K., Marks, J., and Silverman, S., *Late effects of radiation therapy in the head and neck region.* Int J Radiat Oncol Biol Phys, 1995. **31**(5): p. 1141-64.
 42. Eisbruch, A., Kim, H.M., Terrell, J.E., Marsh, L.H., Dawson, L.A., and Ship, J.A., *Xerostomia and its predictors following parotid-sparing*

- irradiation of head-and-neck cancer*. Int J Radiat Oncol Biol Phys, 2001. **50**(3): p. 695-704.
43. van Acker, F., Flamen, P., Lambin, P., Maes, A., Kutcher, G.J., Weltens, C., Hermans, R., Baetens, J., Dupont, P., Rijnders, A., Maes, A., van den Bogaert, W., and Mortelmans, L., *The utility of SPECT in determining the relationship between radiation dose and salivary gland dysfunction after radiotherapy*. Nucl Med Commun, 2001. **22**(2): p. 225-31.
44. Bar-Shalom, R., Valdivia, A.Y., and Blaurock, M.D., *PET imaging in oncology*. Seminars in nuclear medicine, 2000. **30**(3): p. 150-85.
45. Roesink, J.M., Moerland, M.A., Hoekstra, A., Van Rijk, P.P., and Terhaard, C.H., *Scintigraphic assessment of early and late parotid gland function after radiotherapy for head-and-neck cancer: a prospective study of dose-volume response relationships*. Int J Radiat Oncol Biol Phys, 2004. **58**(5): p. 1451-60.
46. Bussels, B., Maes, A., Flamen, P., Lambin, P., Erven, K., Hermans, R., Nuyts, S., Weltens, C., Cecere, S., Lesaffre, E., and Van den Bogaert, W., *Dose-response relationships within the parotid gland after radiotherapy for head and neck cancer*. Radiother Oncol, 2004. **73**(3): p. 297-306.
47. NCI. *Head and Neck Cancer*. 2009.

48. Argiris, A., Karamouzis, M.V., Raben, D., and Ferris, R.L., *Head and neck cancer*. Lancet, 2008. **371**(9625): p. 1695-709.
49. Adelstein, D.J., Li, Y., Adams, G.L., Wagner, H., Jr., Kish, J.A., Ensley, J.F., Schuller, D.E., and Forastiere, A.A., *An intergroup phase III comparison of standard radiation therapy and two schedules of concurrent chemoradiotherapy in patients with unresectable squamous cell head and neck cancer*. J Clin Oncol, 2003. **21**(1): p. 92-8.
50. Bernier, J., Domenge, C., Ozsahin, M., Matuszewska, K., Lefebvre, J.L., Greiner, R.H., Giralt, J., Maingon, P., Rolland, F., Bolla, M., Cognetti, F., Bourhis, J., Kirkpatrick, A., and van Glabbeke, M., *Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer*. N Engl J Med, 2004. **350**(19): p. 1945-52.
51. Kuhnt, T., Becker, A., Bloching, M., Schubert, J., Klautke, G., Fietkau, R., and Dunst, J., *Phase II trial of a simultaneous radiochemotherapy with cisplatin and paclitaxel in combination with hyperfractionated-accelerated radiotherapy in locally advanced head and neck tumors*. Med Oncol, 2006. **23**(3): p. 325-33.
52. Eisbruch, A., Marsh, L.H., Dawson, L.A., Bradford, C.R., Teknos, T.N., Chepeha, D.B., Worden, F.P., Urban, S., Lin, A., Schipper, M.J., and Wolf, G.T., *Recurrences near base of skull after IMRT for head-and-neck*

- cancer: implications for target delineation in high neck and for parotid gland sparing.* Int J Radiat Oncol Biol Phys, 2004. **59**(1): p. 28-42.
53. Garden, A.S., Morrison, W.H., Wong, P.F., Tung, S.S., Rosenthal, D.I., Dong, L., Mason, B., Perkins, G.H., and Ang, K.K., *Disease-control rates following intensity-modulated radiation therapy for small primary oropharyngeal carcinoma.* Int J Radiat Oncol Biol Phys, 2007. **67**(2): p. 438-44.
 54. Bonner, J.A., Harari, P.M., Giralt, J., Azarnia, N., Shin, D.M., Cohen, R.B., Jones, C.U., Sur, R., Raben, D., Jassem, J., Ove, R., Kies, M.S., Baselga, J., Youssoufian, H., Amellal, N., Rowinsky, E.K., and Ang, K.K., *Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck.* N Engl J Med, 2006. **354**(6): p. 567-78.
 55. Denis, F., Garaud, P., Bardet, E., Alfonsi, M., Sire, C., Germain, T., Bergerot, P., Rhein, B., Tortochaux, J., and Calais, G., *Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma.* J Clin Oncol, 2004. **22**(1): p. 69-76.
 56. Colevas, A.D., *Chemotherapy options for patients with metastatic or recurrent squamous cell carcinoma of the head and neck.* J Clin Oncol, 2006. **24**(17): p. 2644-52.

57. Abendstein, H., Nordgren, M., Boysen, M., Jannert, M., Silander, E., Ahlner-Elmqvist, M., Hammerlid, E., and Bjordal, K., *Quality of life and head and neck cancer: a 5 year prospective study*. Laryngoscope, 2005. **115**(12): p. 2183-92.
58. Lavelle, C.L.D., *Saliva: Applied Oral Physiology*. 1988, London: Wright.
59. Chambers, M.S., Toth, B.B., Martin, J.W., Fleming, T.J., and Lemon, J.C., *Oral and dental management of the cancer patient: prevention and treatment of complications*. Support Care Cancer, 1995. **3**(3): p. 168-75.
60. Dawes, C., *Physiological factors affecting salivary flow rate, oral sugar clearance, and the sensation of dry mouth in man*. J Dent Res, 1987. **66 Spec No**: p. 648-53.
61. Percival, R.S., Challacombe, S.J., and Marsh, P.D., *Flow rates of resting whole and stimulated parotid saliva in relation to age and gender*. J Dent Res, 1994. **73**(8): p. 1416-20.
62. Mira, J.G., Wescott, W.B., Starcke, E.N., and Shannon, I.L., *Some factors influencing salivary function when treating with radiotherapy*. Int J Radiat Oncol Biol Phys, 1981. **7**(4): p. 535-41.
63. Dreizen, S., Brown, L.R., Daly, T.E., and Drane, J.B., *Prevention of xerostomia-related dental caries in irradiated cancer patients*. J Dent Res, 1977. **56**(2): p. 99-104.

64. Eneroth, C.M., Henrikson, C.O., and Jakobsson, P.A., *Effect of fractionated radiotherapy on salivary gland function*. Cancer, 1972. **30**(5): p. 1147-53.
65. Franzen, L., Funegard, U., Ericson, T., and Henriksson, R., *Parotid gland function during and following radiotherapy of malignancies in the head and neck. A consecutive study of salivary flow and patient discomfort*. Eur J Cancer, 1992. **28**(2-3): p. 457-62.
66. Stephens, L.C., Schultheiss, T.E., Price, R.E., Ang, K.K., and Peters, L.J., *Radiation apoptosis of serous acinar cells of salivary and lacrimal glands*. Cancer, 1991. **67**(6): p. 1539-43.
67. Stephens, L.C., Schultheiss, T.E., Small, S.M., Ang, K.K., and Peters, L.J., *Response of parotid gland organ culture to radiation*. Radiat Res, 1989. **120**(1): p. 140-53.
68. Paardekooper, G.M., Cammelli, S., Zeilstra, L.J., Coppes, R.P., and Konings, A.W., *Radiation-induced apoptosis in relation to acute impairment of rat salivary gland function*. Int J Radiat Biol, 1998. **73**(6): p. 641-8.
69. Konings, A.W., Coppes, R.P., and Vissink, A., *On the mechanism of salivary gland radiosensitivity*. Int J Radiat Oncol Biol Phys, 2005. **62**(4): p. 1187-94.

70. Coppes, R.P., Roffel, A.F., Zeilstra, L.J., Vissink, A., and Konings, A.W., *Early radiation effects on muscarinic receptor-induced secretory responsiveness of the parotid gland in the freely moving rat*. Radiat Res, 2000. **153**(3): p. 339-46.
71. Konings, A.W., Faber, H., Cotteleer, F., Vissink, A., and Coppes, R.P., *Secondary radiation damage as the main cause for unexpected volume effects: a histopathologic study of the parotid gland*. Int J Radiat Oncol Biol Phys, 2006. **64**(1): p. 98-105.
72. Milano, M.T., Constine, L.S., and Okunieff, P., *Normal tissue tolerance dose metrics for radiation therapy of major organs*. Semin Radiat Oncol, 2007. **17**(2): p. 131-40.
73. Pavy, J.J., Denekamp, J., Letschert, J., Littbrand, B., Mornex, F., Bernier, J., Gonzales-Gonzales, D., Horiot, J.C., Bolla, M., and Bartelink, H., *EORTC Late Effects Working Group. Late Effects toxicity scoring: the SOMA scale*. Int J Radiat Oncol Biol Phys, 1995. **31**(5): p. 1043-7.
74. Pavy, J.J., Denekamp, J., Letschert, J., Littbrand, B., Mornex, F., Bernier, J., Gonzales-Gonzales, D., Horiot, J.C., Bolla, M., and Bartelink, H., *EORTC Late Effects Working Group. Late effects toxicity scoring: the SOMA scale*. Radiother Oncol, 1995. **35**(1): p. 11-5.
75. Rosenthal, D.I., Mendoza, T.R., Chambers, M.S., Burkett, V.S., Garden, A.S., Hessel, A.C., Lewin, J.S., Ang, K.K., Kies, M.S., Gning, I., Wang,

- X.S., and Cleeland, C.S., *The M. D. Anderson symptom inventory-head and neck module, a patient-reported outcome instrument, accurately predicts the severity of radiation-induced mucositis*. Int J Radiat Oncol Biol Phys, 2008. **72**(5): p. 1355-61.
76. Meirovitz, A., Murdoch-Kinch, C.A., Schipper, M., Pan, C., and Eisbruch, A., *Grading xerostomia by physicians or by patients after intensity-modulated radiotherapy of head-and-neck cancer*. Int J Radiat Oncol Biol Phys, 2006. **66**(2): p. 445-53.
 77. Ship, J.A., Patton, L.L., and Tylena, C.A., *An assessment of salivary function in healthy premenopausal and postmenopausal females*. J Gerontol, 1991. **46**(1): p. M11-5.
 78. Lee, J.R., Madsen, M.T., Bushnell, D., and Menda, Y., *A threshold method to improve standardized uptake value reproducibility*. Nucl Med Commun, 2000. **21**(7): p. 685-90.
 79. Graham, M.M., Peterson, L.M., and Hayward, R.M., *Comparison of simplified quantitative analyses of FDG uptake*. Nuclear medicine and biology, 2000. **27**(7): p. 647-55.
 80. Wang, H., Dong, L., Lii, M.F., Lee, A.L., de Crevoisier, R., Mohan, R., Cox, J.D., Kuban, D.A., and Cheung, R., *Implementation and validation of a three-dimensional deformable registration algorithm for targeted*

- prostate cancer radiotherapy*. Int J Radiat Oncol Biol Phys, 2005. **61**(3): p. 725-35.
81. Wang, H., Dong, L., O'Daniel, J., Mohan, R., Garden, A.S., Ang, K.K., Kuban, D.A., Bonnen, M., Chang, J.Y., and Cheung, R., *Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy*. Phys Med Biol, 2005. **50**(12): p. 2887-905.
 82. Thirion, J.P., *Image matching as a diffusion process: an analogy with Maxwell's demons*. Med Image Anal, 1998. **2**(3): p. 243-60.
 83. Court, L.E. and Dong, L., *Automatic registration of the prostate for computed-tomography-guided radiotherapy*. Medical physics, 2003. **30**(10): p. 2750-7.
 84. Zhang, L., Dong, L., Court, L., *Validation of CT-Assisted Targeting (CAT) software for soft tissue and bony target localization*, in *The American Association of Physicists in Medicine (AAPM) 47th Annual Meeting*. 2005: Seattle, WA.
 85. Zhang, L., Garden, A.S., Lo, J., Ang, K.K., Ahamad, A., Morrison, W.H., Rosenthal, D.I., Chambers, M.S., Zhu, X.R., Mohan, R., and Dong, L., *Multiple regions-of-interest analysis of setup uncertainties for head-and-neck cancer radiotherapy*. Int J Radiat Oncol Biol Phys, 2006. **64**(5): p. 1559-69.

86. Pan, T., Liu, H., Xuamin, W., Mawlawi, O., and Luo, D., *Using SUV in PET Images to Delineate Tumors in Phillips Pinnacle Treatment Planning System*. 2007, The University of Texas MD Anderson Cancer Center: Houston. p. 4.
87. Kim, C.K. and Gupta, N.C., *Dependency of standardized uptake values of fluorine-18 fluorodeoxyglucose on body size: comparison of body surface area correction and lean body mass correction*. Nucl Med Commun, 1996. **17**(10): p. 890-4.
88. Hanley, J.A. and McNeil, B.J., *A method of comparing the areas under receiver operating characteristic curves derived from the same cases*. Radiology, 1983. **148**(3): p. 839-43.
89. Kunkel, M., Forster, G.J., Reichert, T.E., Kutzner, J., Benz, P., Bartenstein, P., and Wagner, W., *Radiation response non-invasively imaged by [18F]FDG-PET predicts local tumor control and survival in advanced oral squamous cell carcinoma*. Oral Oncol, 2003. **39**(2): p. 170-7.
90. Pan, T. and Mawlawi, O., *PET/CT in radiation oncology*. Med Phys, 2008. **35**(11): p. 4955-66.
91. Weber, W.A., *Use of PET for monitoring cancer therapy and for predicting outcome*. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2005. **46**(6): p. 983-95.

92. Weber, W.A., *Positron emission tomography as an imaging biomarker*. J Clin Oncol, 2006. **24**(20): p. 3282-92.
93. Weber, W.A. and Figlin, R., *Monitoring cancer treatment with PET/CT: does it make a difference?* Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 2007. **48**: p. 36S-44S.
94. Raylman, R.R., Fisher, S.J., Brown, R.S., Ethier, S.P., and Wahl, R.L., *Fluorine-18-fluorodeoxyglucose-guided breast cancer surgery with a positron-sensitive probe: validation in preclinical studies*. J Nucl Med, 1995. **36**(10): p. 1869-74.
95. Wahl, R.L., Jacene, H., Kasamon, Y., and Lodge, M.A., *From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors*. J Nucl Med, 2009. **50 Suppl 1**: p. 122S-50S.
96. Soret, M., Bacharach, S.L., and Buvat, I., *Partial-volume effect in PET tumor imaging*. J Nucl Med, 2007. **48**(6): p. 932-45.
97. Beyer, T., Tellmann, L., Nickel, I., and Pietrzyk, U., *On the use of positioning aids to reduce misregistration in the head and neck in whole-body PET/CT studies*. J Nucl Med, 2005. **46**(4): p. 596-602.
98. Brown, R.S., Fisher, S.J., and Wahl, R.L., *Autoradiographic evaluation of the intra-tumoral distribution of 2-deoxy-D-glucose and monoclonal*

- antibodies in xenografts of human ovarian adenocarcinoma.* J Nucl Med, 1993. **34**(1): p. 75-82.
99. Brown, R.S., Leung, J.Y., Kison, P.V., Zasadny, K.R., Flint, A., and Wahl, R.L., *Glucose transporters and FDG uptake in untreated primary human non-small cell lung cancer.* J Nucl Med, 1999. **40**(4): p. 556-65.
 100. Brown, R.S. and Wahl, R.L., *Overexpression of Glut-1 glucose transporter in human breast cancer. An immunohistochemical study.* Cancer, 1993. **72**(10): p. 2979-85.
 101. Mawlawi, O., Pan, T., and Macapinlac, H.A., *PET/CT imaging techniques, considerations, and artifacts.* J Thorac Imaging, 2006. **21**(2): p. 99-110.
 102. Britz-Cunningham, S.H., Millstine, J.W., and Gerbaudo, V.H., *Improved discrimination of benign and malignant lesions on FDG PET/CT, using comparative activity ratios to brain, basal ganglia, or cerebellum.* Clin Nucl Med, 2008. **33**(10): p. 681-7.
 103. Hylton, N., *Dynamic contrast-enhanced magnetic resonance imaging as an imaging biomarker.* J Clin Oncol, 2006. **24**(20): p. 3293-8.
 104. Winter, P.M., Morawski, A.M., Caruthers, S.D., Fuhrhop, R.W., Zhang, H., Williams, T.A., Allen, J.S., Lacy, E.K., Robertson, J.D., Lanza, G.M., and Wickline, S.A., *Molecular imaging of angiogenesis in early-stage*

- atherosclerosis with alpha(v)beta3-integrin-targeted nanoparticles.* Circulation, 2003. **108**(18): p. 2270-4.
105. Dornfeld, K., Hopkins, S., Simmons, J., Spitz, D.R., Menda, Y., Graham, M., Smith, R., Funk, G., Karnell, L., Karnell, M., Dornfeld, M., Yao, M., and Buatti, J., *Posttreatment FDG-PET uptake in the supraglottic and glottic larynx correlates with decreased quality of life after chemoradiotherapy.* Int J Radiat Oncol Biol Phys, 2008. **71**(2): p. 386-92.
 106. Guerrero, T., Johnson, V., Hart, J., Pan, T., Khan, M., Luo, D., Liao, Z., Ajani, J., Stevens, C., and Komaki, R., *Radiation pneumonitis: local dose versus [18F]-fluorodeoxyglucose uptake response in irradiated lung.* Int J Radiat Oncol Biol Phys, 2007. **68**(4): p. 1030-5.
 107. Jeraj, R., Cao, Y., Ten Haken, R.K., Hahn, C., and Marks, L., *Imaging for assessment of radiation-induced normal tissue effects.* Int J Radiat Oncol Biol Phys. **76**(3 Suppl): p. S140-4.
 108. Jensen, T., Jacques, L., Caplan, S., Tillman, K., and Roche, J., *Decision Memorandum for Positron Emission Tomography (FDG) for Solid Tumors and Myeloma,* in CAG-00181R. 2009, Centers for Medicare and Medicaid.
 109. Davies, A.N., Broadley, K., and Beighton, D., *Xerostomia in patients with advanced cancer.* J Pain Symptom Manage, 2001. **22**(4): p. 820-5.

110. van de Water, T.A., Bijl, H.P., Westerlaan, H.E., and Langendijk, J.A., *Delineation guidelines for organs at risk involved in radiation-induced salivary dysfunction and xerostomia*. Radiother Oncol, 2009. **93**(3): p. 545-52.
111. Harnsberger, H.R., *Diagnostic and surgical imaging anatomy: brain, head and neck, spine*. Head & neck, 2006.
112. O'Daniel, J.C., Rosenthal, D.I., Garden, A.S., Barker, J.L., Ahamad, A., Ang, K.K., Asper, J.A., Blanco, A.I., de Crevoisier, R., Holsinger, F.C., Patel, C.B., Schwartz, D.L., Wang, H., and Dong, L., *The effect of dental artifacts, contrast media, and experience on interobserver contouring variations in head and neck anatomy*. Am J Clin Oncol, 2007. **30**(2): p. 191-8.
113. Erdi, Y.E., Mawlawi, O., Larson, S.M., Imbriaco, M., Yeung, H., Finn, R., and Humm, J.L., *Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding*. Cancer, 1997. **80**(12): p. 2505-9.
114. Fischer, D. and Ship, J.A., *Effect of age on variability of parotid salivary gland flow rates over time*. Age Ageing, 1999. **28**(6): p. 557-61.
115. Ship, J.A., Nolan, N.E., and Puckett, S.A., *Longitudinal analysis of parotid and submandibular salivary flow rates in healthy, different-aged adults*. J Gerontol A Biol Sci Med Sci, 1995. **50**(5): p. M285-9.

116. Tucker, S.L., Cheung, R., Dong, L., Liu, H.H., Thames, H.D., Huang, E.H., Kuban, D., and Mohan, R., *Dose-volume response analyses of late rectal bleeding after radiotherapy for prostate cancer*. Int J Radiat Oncol Biol Phys, 2004. **59**(2): p. 353-65.
117. Burman, C., Kutcher, G.J., Emami, B., and Goitein, M., *Fitting of normal tissue tolerance data to an analytic function*. Int J Radiat Oncol Biol Phys, 1991. **21**(1): p. 123-35.
118. Royston, P. and Altman, D.G., *Regression Using Fractional Polynomials of Continuous Covariates - Parsimonious Parametric Modeling*. Applied Statistics-Journal of the Royal Statistical Society Series C, 1994. **43**(3): p. 429-467.
119. Sauerbrei, W. and Royston, P., *Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials*. Journal of the Royal Statistical Society Series a-Statistics in Society, 1999. **162**: p. 71-94.
120. Schwartz-Arad, D., Michaeli, Y., and Zajicek, G., *Compensatory hyperplasia of the rat submandibular gland following unilateral extirpation*. J Dent Res, 1991. **70**(10): p. 1328-31.

121. Dogan, N., Siebers, J.V., Keall, P.J., Lerma, F., Wu, Y., Fatyga, M., Williamson, J.F., and Schmidt-Ullrich, R.K., *Improving IMRT dose accuracy via deliverable Monte Carlo optimization for the treatment of head and neck cancer patients*. Med Phys, 2006. **33**(11): p. 4033-43.

Vita

Blake Allen Cannon was born in Decatur, Georgia on February 21, 1981, the son of Ken Allen Cannon and Vicki Avery Cannon. After completing his work at Dacula High School, Dacula, Georgia in 1999, he entered the University of Georgia, Athens, Georgia. His interest in medical physics began while working in the EPA's National Ultraviolet Monitoring Center, Athens, Georgia as a National Science Foundation Research Experience for Undergraduates (REU) Summer Fellow. Blake graduated from UGA with a B.S. in Physics and Astronomy, cum laude. He married Cara Megan Williams, the daughter of Richard F. Williams and Pamela J. Stevens on June 21, 2003. In September of 2003 he entered The University of Texas Health Science Center at Houston Graduate School of Biomedical Sciences. Working under Jingfei Ma, Ph.D., he obtained his M.S. degree in 2006, specializing in quantitative magnetic resonance imaging. However, his proudest achievement was the birth of his daughter Avery Anne Cannon on March 30, 2006. In September of 2006, Blake began work on his Ph.D. under the supervision of Lei Dong, Ph.D.

Permanent address:

7575 Cambridge Street #904

Houston, TX 77054

Blake.cannon@gmail.com

twinsparadox@hotmail.com